目　录

第一章　内积空间与矩阵的Jordan标准形

在《线性代数》中，已经学过了向量、矩阵、线性空间以及线性变换的基本知识。作为与本课程的衔接，我们首先对线性空间和线性变换作一简要概述。

第一节　线性空间及线性变换概述

一、线性空间

定义1　设V是一个非空集合，F是一个数域，如果V有下面两种运算：

（1）加法运算：对V中任意两个元素X、Y，都有V中唯一确定的一个元素Z与之对应，称Z为X与Y的和，记作$Z=X+Y$；

（2）数乘运算：对F中任一数k及V中任一元素X，都有V中唯一确定的一个元素r与之对应，称r为k与X的积，记作$r=kX$

并且这两种运算满足以下八条运算规律：

设X，Y，$Z\in V$；k，$\lambda\in F$

$1°\ (X+Y)+Z=X+(Y+Z)$　　（结合律）

$2°\ X+Y=Y+X$　　（交换律）

$3°$ 在V中存在一个元素0，对V中任一元素X，都有

$X+0=X$　　（零元律）

元素0称为V的零元素。

4°对V中每一个元素X，都有V中元素X'存在，使得

$X+X'=0$ （负元律）

X'称为元素X的负元素，记作$X'=-X$

5° $1\cdot X=X$ $(1\in F)$ （单位元律）

6° $(k\lambda)X=k(\lambda X)=\lambda(kX)$ （数乘结合律）

7° $(k+\lambda)X=kX+\lambda X$ （数量加法分配律）

8° $k(X+Y)=kX+kY$ （向量加法分配律）

则称V是数域F上的线性空间。

如果F为实数域(即$F=R$)，则称V为实线性空间；若F为复数域(即$F=C$)，则称V为复线性空间。

从定义可直接得出线性空间的一些基本性质：

性质1 零元素是唯一的。

性质2 负元素是唯一的。

性质3 $0\cdot X=0$ $(0\in F;\ 0\in V)$

$k\cdot 0=0$ $(k\in F;\ 0\in V)$

$(-1)X=-X$ $(-1\in F;\ X\in V)$

性质4 如果$kX=0$，则

$k=0$或者$X=0$ $(k\in F;\ 0\in V)$

定义2 如果线性空间V中有n个线性无关的元素，且任意$n+1$个元素都是线性相关的，则称n为线性空间V的维数，记作$\dim(V)=n$；维数是n的线性空间，称为n维线性空间，常记为V_n。

定义3 设e_1，e_2，e_3，…，e_n是n维线性空间V中的几个线性无关的元素，则称e_1，e_2，…，e_n是线性空间V的一组基底。此时V中任一元素a，均可由这组基底唯一地线性表出。如

$$a=a_1e_1+a_2e_2+\cdots+a_ne_n$$

其中系数$(a_1, a_2, \cdots, a_n)$称为元素a在$e_1, e_2, \cdots, e_n$这组基底下的坐标。

建立了坐标以后，就可以把抽象的元素与具体的数组向量$(a_1, a_2, \cdots, a_n)$联系起来了。因此，我们把V中的元素a，也叫做向量。此时线性空间V又叫做向量空间。

注意，线性空间中同一个向量在不同基底下的坐标是不相同的。

一般地，基底变换与坐标变换的关系如下：

定理1 设$e_1, e_2, \cdots, e_n$与$e_1', e_2', \cdots, e_n'$为n维线性空间V中的两组基底，且

$$\begin{aligned} e_1' &= a_{11}e_1 + a_{21}e_2 + \cdots + a_{n1}e_n \\ e_2' &= a_{12}e_1 + a_{22}e_2 + \cdots + a_{n2}e_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ e_n' &= a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n \end{aligned}$$

或

$$\begin{pmatrix} e_1' \\ e_2' \\ \vdots \\ e_n' \end{pmatrix} = A^{\mathrm{T}} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \quad 其中A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

如果V中任意向量a，在此二组基底下的坐标分别是$X=(x_1, x_2, \cdots, x_n)$和$X'=(x_1', x_2', \cdots, x_n')$则

$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = A^{-1} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

证明 由于$e_1, e_2, \cdots, e_n$线性无关，则矩阵A是满秩的，因而A^{-1}存在。

因为 $\alpha=\sum_{j=1}^{n} x_j e_j=\sum_{i=1}^{n} x_i' e_i'=\sum_{i=1}^{n} x_i'\left(\sum_{j=1}^{n} a_{ji} e_j\right)$

$$=\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ji} x_i'\right) e_j$$

则 $x_j=\sum_{i=1}^{n} a_{ji} x_i' \quad (j=1,\ 2,\ \cdots,\ n)$

故
$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}=A^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

定理的逆也成立。即如果

$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}=A^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$ 则有

$$\begin{pmatrix} e_1' \\ e_2' \\ \vdots \\ e_n' \end{pmatrix}=A^{\mathrm{T}}\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

矩阵A称为由基底e_1，e_2，…，e_n到基底e_1'，e_2'，…，e_n'的**过渡矩阵**。

利用定理1就可由基底变换得到坐标交换；反过来，也可由坐标变换求得基底变换。

定义4 设V是数域F上的一个线性空间，S是V中的一个非空子集，若按V中的“加法”和“数乘”两种运算，S也构

成一个线性空间，则称S是V的一个线性子空间（或称向量子空间），简称子空间。

由子空间定义，显然，在任意线性空间中，由单个的零元素所组成的集合为它的一个子空间，叫做零子空间。线性空间本身也为它自己的一个子空间。这两个子空间称为平凡子空间，其它的子空间称为非平凡子空间。

线性空间V的非空子集构成V的子空间的充要条件如下：

定理2 线性空间V的非空子集S构成V的子空间的充要条件是：如果X，$Y\in S$；k，$\lambda\in F$，则$kX+\lambda Y\in S$

证明 若S是V的子空间，由子空间定义4，S对V中的两种运算（"加法"和"数乘"）是封闭的。即如果$X,Y\in S$；k，$\lambda\in F$，则$kX+\lambda Y\in S$

反之，若X，$Y\in S$，且当k，$\lambda\in F$时，有$kX+\lambda Y\in S$即S对V中的两种运算是封闭的，则有$0=0\cdot X+0\cdot Y\in S$，$(-1)X=-X\in S$，故S满足线性空间定义1中的3°、4°；又因S是V的一部分，故V中的运算对S来说，规律1°、2°、5°、6°、7°、8°显然是满足的，所以S构成V的子空间。

二、线性变换

若在线性空间V中存在某种法则$\mathscr{A}$，使V中任一向量X对应于V中唯一向量Y。则称此法则$\mathscr{A}$为V的一个变换，记作$\mathscr{A}(X)=Y$。Y称为X的象，X叫做Y的象源。

下面介绍的线性变换是最简单、最基本的一种变换。

定义5 设$\mathscr{A}$是线性空间V上的一个变换，如满足下列条件：

(1) $\mathscr{A}(X+Y)=\mathscr{A}(X)+\mathscr{A}(Y)$ $(X,\ Y\in V)$

(2) $\mathscr{A}(kX)=k\mathscr{A}(X)$ $(X\in V;\ k\in F)$

则称$\mathscr{A}$是V上的一个线性变换(或称线性算子)。

根据上面的定义可得出：一个变换$\mathscr{A}$为线性变换的充要条件是：对线性空间V中任意两个向量X，Y及数域F中任意二数k，λ，使得

$$\mathscr{A}(kX+\lambda Y)=k\mathscr{A}(X)+\lambda\mathscr{A}(Y)$$

成立。

由上述结论可以证明在线性空间V中，下列变换都是V上的线性变换。

恒等变换：$\mathscr{A}(X)=X$　又称单位变换

可记为　$\mathscr{E}(X)=X$

数乘变换：$\mathscr{A}(X)=kX$　　$(k\in F)$

零变换：$\theta(X)=0$　　$(0\in V)$

设$\mathscr{A}$是线性空间V上的线性变换，从定义5可直接推出如下简单性质：

性质1　$\mathscr{A}(0)=0$；$\mathscr{A}(-X)=-\mathscr{A}(X)$

即线性变换把零元仍变为零元；把X的负元$-X$变为X的象$\mathscr{A}(X)$的负元。

性质2　设 $Y=k_1X_1+k_2X_2+\cdots+k_rX_r$

$$(k_1,\ k_2,\ \cdots,\ k_r\in F)$$

则　$\mathscr{A}(Y)=k_1\mathscr{A}(X_1)+k_2\mathscr{A}(X_2)+\cdots+k_r\mathscr{A}(k_r)$

即线性变换保持线性组合与线性关系式不变。

性质3　若X_1，X_2，$\cdots$，X_r线性相关，则$\mathscr{A}(X_1)$，$\mathscr{A}(X_2)$，$\cdots$，$\mathscr{A}(X_r)$也线性相关。但其逆一般不成立，即线性变换$\mathscr{A}$可能会把线性无关向量组变成线性相关的。例如零变换就是这样。

定义6　设$\mathscr{A}$，$\mathscr{B}$是线性空间V上的两个线性变换，若对V中每一个向量X，都有

$$\mathscr{A}(X)=\mathscr{B}(X)$$

则称$\mathscr{A}$和$\mathscr{B}$相等，即$\mathscr{A}=\mathscr{B}$

定义7 设$\mathscr{A}$，$\mathscr{B}$是线性空间V上的两个线性变换，X是V中任意一个向量，k是数域F中一个数，定义：

加法：　$(\mathscr{A}+\mathscr{B})(X)=\mathscr{A}(X)+\mathscr{B}(X)$

数乘法：$(k\mathscr{A})(X)=k\mathscr{A}(X)$

乘积：　$(\mathscr{A}\mathscr{B})(X)=\mathscr{A}(\mathscr{B}(X))$

由定义5及定义7可得如下定理

定理3 设$\mathscr{A}$和$\mathscr{B}$是线性空间V上的两个线性变换，则

(1) $\mathscr{A}+\mathscr{B}$也是线性变换；

(2) $k\mathscr{A}$也是线性变换；

(3) $\mathscr{A}\mathscr{B}$也是线性变换。

注意：两个线性变换的乘积一般不满足交换律，即

$$\mathscr{A}\mathscr{B}\neq\mathscr{B}\mathscr{A}$$

例如，设$\mathscr{A}(f(x))=\dfrac{df}{dx}$，$\mathscr{B}(f(x))=\int_0^x f(x)dx$

$$\mathscr{A}\mathscr{B}(f(x))=f(x)$$

但　　$\mathscr{B}\mathscr{A}(f(x))=f(x)-f(0)$ （c为任意常数）

因此　$\mathscr{A}\mathscr{B}\neq\mathscr{B}\mathscr{A}$

定义8 设$\mathscr{A}$和$\mathscr{B}$是线性空间V上的两个线性变换，若满足

$$\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}=\mathscr{E}\quad（\mathscr{E}\text{是}V\text{上的单位变换}）$$

则称线性变换$\mathscr{B}$是$\mathscr{A}$的逆变换，记作$\mathscr{A}^{-1}=\mathscr{B}$

线性变换可以通过向量的坐标来建立它与矩阵的关系。

设V是数域F上的一个n维线性空间，e_1，e_2，…，e_n是V的一组基底，$\mathscr{A}$是V上的一个线性变换，则对V中任意一个向量a，有

$$a=a_1e_1+a_2e_2+\cdots+a_ne_n$$

对上式两端取线性变换$\mathscr{A}$，得

$$\mathscr{A}(a)=a_1\mathscr{A}(e_1)+a_2\mathscr{A}(e_2)+\cdots+a_n\mathscr{A}(e_n)$$

若
$$\begin{cases}\mathscr{A}(e_1)=a_{11}e_1+a_{21}e_2+\cdots+a_{n1}e_n\\ \mathscr{A}(e_2)=a_{12}e_1+a_{22}e_2+\cdots+a_{n2}e_n\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \mathscr{A}(e_n)=a_{1n}e_1+a_{2n}e_2+\cdots+a_{nn}e_n\end{cases}$$

用矩阵表示，可写为

$$\begin{pmatrix}\mathscr{A}(e_1)\\ \mathscr{A}(e_2)\\ \vdots\\ \mathscr{A}(e_n)\end{pmatrix}=A^{\mathrm{T}}\begin{pmatrix}e_1\\ e_2\\ \vdots\\ e_n\end{pmatrix}$$

其中 $$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$$

对上式两边转置，得

$$\mathscr{A}[e_1,\ e_2,\ \cdots,\ e_n]=[e_1,\ e_2,\ \cdots,\ e_n]A$$

矩阵A称为线性变换$\mathscr{A}$在基底e_1，e_2，…，e_n下的矩阵。

再进一步讨论在线性空间中，任一向量X与它的象$\mathscr{A}(X)$在同一基底下坐标之间的关系。

定理4 设$\mathscr{A}$是n维线性空间V中的一个线性变换，它在基底e_1，e_2，…，e_n下的矩阵为A。若向量X与$\mathscr{A}(X)$在基底e_1，e_2，…，e_n下的坐标分别为$(x_1,\ x_2,\ \cdots,\ x_n)$和$(y_1,\ y_2,\ \cdots,\ y_n)$，则坐标间的关系为式：

$$Y=AX$$

其中　$X^T=[x_1, x_2, \cdots, x_n]$，$Y^T=[y_1, y_2, \cdots, y_n]$

证明　因为

$$X=\sum_{i=1}^{n} x_i e_i$$

对上式两边取变换，得

$$\mathscr{A}(X)=\sum_{i=1}^{n} x_i \mathscr{A}(e_i)$$

设

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则基向量的线性变换公式为：

$$\mathscr{A}(e_i)=\sum_{j=1}^{n} a_{ji} e_j \quad (i=1, 2, \cdots, n)$$

于是，有

$$\mathscr{A}(X)=\sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} a_{ji} e_j\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ji} x_i\right) e_j$$

所以　$y_j=\sum_{i=1}^{n} a_{ji} x_i \quad (j=1, 2, \cdots, n)$

即得　$Y=AX$

线性变换的矩阵是与空间中的一组基底联系在一起的。一般说来，随着基底的改变，同一个线性变换就有不同的矩阵。

线性变换的矩阵是如何随着基底的变化而改变的呢？

定理5　设$e_1, e_2, \cdots, e_n$与$e_1', e_2', \cdots, e_n'$为n维线

性空间V中的两组基底。线性变换$\mathscr{A}$在此两组基底下的矩阵分别为$A=[a_{ij}]$与$B=[b_{ij}]$；又由基底e_1，e_2，…，e_n到基底e_1'，e_2'，…，e_n'的过渡矩阵为$C=[c_{ij}]$，则矩阵A、B、C之间满足如下关系式：

$$B=C^{-1}AC$$

证明　由假设知

$$e_i'=\sum_{j=1}^{n}c_{ji}e_j \qquad (i=1,\ 2,\ \cdots,\ n)$$

$$\mathscr{A}(e_i)=\sum_{j=1}^{n}a_{ji}e_j \qquad (i=1,\ 2,\ \cdots,\ n)$$

$$\mathscr{A}(e_i')=\sum_{j=1}^{n}b_{ji}e_j' \qquad (i=1,\ 2,\ \cdots,\ n)$$

于是有

$$\mathscr{A}(e_i')=\mathscr{A}\left(\sum_{j=1}^{n}c_{ji}e_j\right)=\sum_{j=1}^{n}c_{ji}\mathscr{A}(e_j)$$

$$=\sum_{j=1}^{n}c_{ji}\left(\sum_{k=1}^{n}a_{kj}e_k\right)=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}a_{kj}c_{ji}\right)e_k$$

$$(i=1,\ 2,\ \cdots,\ n)$$

另一方面有

$$\mathscr{A}(e_i')=\sum_{j=1}^{n}b_{ji}e_j'=\sum_{j=1}^{n}b_{ji}\left(\sum_{k=1}^{n}c_{kj}e_k\right)$$

$$=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}c_{kj}b_{ji}\right)e_k \quad (i=1,\ 2,\ \cdots,\ n)$$

对上两式右边进行比较，得

$$\sum_{j=1}^{n} a_{kj}c_{ji}=\sum_{j=1}^{n} c_{kj}b_{ji} \quad (k,\ i=1,\ 2,\ \cdots,\ n)$$

于是有　$AC=CB$

因为矩阵C满秩，故存在C^{-1}，将其左乘上式两端，即得

$$B=C^{-1}AC$$

例如：设e_1，e_2，e_3为三维线性空间V_3中的一组基底，线性变换$\mathscr{A}$在基底e_1，e_2，e_3下的矩阵为

$$A=\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \\ 2 & 1 & 5 \end{pmatrix}$$

若V_3中另一组基底e'_1，e'_2，e'_3与基底e_1，e_2，e_3有如下关系式：

$$e'_1=e_1 \qquad e'_2=e_1+e_2 \qquad e'_3=e_1+e_2+e_3$$

求线性变换$\mathscr{A}$在e'_1，e'_2，e'_3下的矩阵B。

解　由两组基底的关系式知，过渡矩阵C为

$$C=\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

则

$$C^{-1}=\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

由定理 5 得

$$B=C^{-1}AC=\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & 3 \\ 2 & 1 & 5 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 2 & 4 & 4 \\ -3 & -4 & -6 \\ 2 & 3 & 8 \end{pmatrix}$$

注意：在定理5中，矩阵A，B，C之间的关系是很重要的，有如下的定义

定义9 设A，B为数域F上的两个n阶方阵，若存在可逆的n阶方阵P，使得

$$B=P^{-1}AP$$

则称B是A的相似矩阵，或者说A与B相似，记作$A\sim B$

对A进行$P^{-1}AP$的运算，称为对A进行相似变换，同时称P为把A变成B的**相似变换矩阵**

矩阵的相似关系，具有下面三个性质：

性质1 （反身性） $A\sim A$

性质2 （对称性） 若$A\sim B$，则$B\sim A$

性质3 （传递性） 若$A\sim B$；$B\sim C$，则$A\sim C$

由定义9，可得如下定理：

定理6 线性变换在不同基底下所对应的矩阵是相似的；反之，若两个矩阵相似，则它们可视为同一线性变换在两组不同基底下的对应矩阵。

证明 此定理的前面部分，已由定理5证明了。现在证明后面部分。若$A\sim B$，则矩阵A，B是同一线性变换$\mathscr{A}$在两组不同基底下所对应的矩阵。

设$A=[a_{ij}]$是$\mathscr{A}$在基底e_1，e_2，…，e_n下的矩阵，即有

$$\mathscr{A}(e_i)=\sum_{j=1}^{n}a_{ji}e_j \qquad (i=1,\ 2,\ \cdots,\ n)$$

因为$A\sim B$，则存在可逆方阵$P=[p_{ij}]$，使得

$$B=P^{-1}AP$$

令
$$e_i'=\sum_{j=1}^{n}p_{ji}e_j \qquad (i=1,\ 2,\ \cdots,\ n)$$

显然e_1'，e_2'，…，e_n'也是一组基底。对上式两端取线性变换$\mathscr{A}$，则有

$$\mathscr{A}(e_i')=\mathscr{A}\left(\sum_{j=1}^{n} p_{ji}e_j\right)=\sum_{j=1}^{n} p_{ji}\mathscr{A}(e_j)$$

$$=\sum_{j=1}^{n} p_{ji}\left(\sum_{k=1}^{n} a_{kj}e_k\right)=\sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj}p_{ji}\right)e_k$$

又因　$PB=AP$，若记$B=[b_{ij}]$，则有

$$\sum_{j=1}^{n} a_{kj}p_{ji}=\sum_{j=1}^{n} p_{kj}b_{ji}$$

于是

$$\mathscr{A}(e_i')=\sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{kj}p_{ji}\right)e_k=\sum_{k=1}^{n}\left(\sum_{j=1}^{n} p_{kj}b_{ji}\right)e_k$$

$$=\sum_{j=1}^{n} b_{ji}\left(\sum_{k=1}^{n} p_{kj}e_k\right)=\sum_{j=1}^{n} b_{ji}e_j'$$

所以$B=[b_{ij}]$是变换$\mathscr{A}$在基底e_1'，e_2'，…，e_n'下的矩阵。

相似矩阵有下面的运算法则：

法则1　和的相似矩阵等于相似矩阵的和。

即　$P^{-1}(A_1+A_2)P=P^{-1}A_1P+P^{-1}A^2P$

显然，可推广到

$$P^{-1}(A_1+A_2+\cdots+A_m)P=P^{-1}A_1P+P^{-1}A_2P+\cdots+P^{-1}A_mP$$

法则2　与数k乘积的相似矩阵等于相似矩阵乘以数k，即

$$P^{-1}(kA)P=kP^{-1}AP \qquad (k\in F)$$

法则3　乘积的相似矩阵等于相似矩阵的乘积。即

$$P^{-1}(A_1A_2)P=(P^{-1}A_1P)(P^{-1}A_2P)$$

显然可推广到

$$P^{-1}(A_1A_2\cdots A_m)P=(P^{-1}A_1P)(P^{-1}A_2P)\cdots(P^{-1}A_mP)$$

由此当$A_1=A_2=\cdots=A_m=A$时，有

$$P^{-1}A^mP=(P^{-1}AP)^m$$

以上三个法则利用矩阵运算及逆矩阵定义很容易证明。又由法则1和法则2可以推出相似变换是一个线性变换。

第二节　内积空间

在线性空间中，向量之间只有加法和数乘运算即线性运算，而没有象解析几何中向量的长度、两向量的夹角等度量概念。这些度量概念在很多问题中又十分重要。因此，有必要在线性空间中引进某种度量概念。

鉴于解析几何中，向量的长度与夹角都可以通过向量的内积来表示。据此，我们先引进内积概念，从而引出了Euclid空间。

一、Euclid空间的基本概念

定义1　设V是实数域R上的线性空间，若V中任意二向量X，Y都按某一确定法则对应于唯一确定的实数，记作(X, Y)，它满足下列条件：

(1)　$(X, Y)=(Y, X)$

(2)　$(kX, Y)=k(X, Y)$　　$(k\in R)$

(3)　$(X+Y, Z)=(X, Z)+(Y, Z)$

(4)　$(X, X)\geqslant 0$，当且仅当$X=0$时，$(X, X)=0$则称(X, Y)为向量X与Y的内积。定义了内积的实线性空间V，叫做实内积空间，亦称**Euclid空间**(简称**欧氏空间**)。

〔**例1**〕 在实空间R^n中，对向量$X=(x_1, x_2, \cdots, x_n)$和$Y=(y_1, y_2, \cdots, y_n)$规定

$$(X, Y)=x_1y_1+x_2y_2+\cdots+x_ny_n$$

显然它满足定义1中的四个条件，所以(X, Y)是内积。这样，R^n就成为欧氏空间。

〔**例2**〕 设V是定义在闭区间$[a,b]$上的实值连续函数的全体，则V对于通常函数的加法及数乘运算构成线性空间。若对V中任意二函数$f(x)$，$g(x)$规定

$$(f(x), g(x))=\int_a^b f(x)g(x)\mathrm{d}x$$

则利用积分性质，容易证明$(f(x), g(x))$是内积，从而V就成为欧氏空间。

〔**例3**〕 设$R^{n\times n}$是元素为实数的n阶矩阵的全体，则$R^{n\times n}$对于矩阵的加法及数乘运算构成线性空间。若对$R^{n\times n}$中任意二矩阵$A=[a_{ij}]$，$B=[b_{ij}]$规定

$$(A, B)=\mathrm{tr}(AB^{\mathrm{T}})=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}b_{ij}$$

则易证(A, B)是内积，因而，$R^{n\times n}$也就成为欧氏空间。

从欧氏空间V的定义，可推出内积的下列性质：

性质1 $(0, X)=(X, 0)=0$

因为 $(0, X)=(0\cdot X, X)=0(X, X)=0$

性质2 $(X, kY)=k(X, Y)$

事实上 $(X, kY)=(kY, X)=k(Y, X)=k(X, Y)$

性质3 $(X, Y+Z)=(X, Y)+(X, Z)$

事实上 $(X, Y+Z)=(Y+Z, X)=(Y, X)+(Z, X)$

$$=(X, Y)+(X, Z)$$

性质4 对V中任意向量X_1, X_2, …, X_m; Y_1, Y_2, …, Y_n及实数k_1, k_2, …, k_m; λ_1, λ_2, …, λ_n有

$$\left(\sum_{i=1}^{m} k_i X_i,\ \sum_{j=1}^{n} \lambda_j Y_j\right)=\sum_{i=1}^{m}\sum_{j=1}^{n} k_i \lambda_j (X_i,\ Y_j)$$

事实上 首先用数学归纳法不难将定义中的第(3)条及上述性质3推广到任意有限多项的情形：

$$\left(\sum_{i=1}^{m} X_i,\ Y\right)=\sum_{i=1}^{m}(X_i,\ Y)$$

及 $$\left(X,\ \sum_{j=1}^{n} Y_j\right)=\sum_{j=1}^{n}(X,\ Y_j)$$

然后，再利用定义中的第(2)条及上述性质2即可推得

$$\left(\sum_{i=1}^{m} k_i X_i,\ \sum_{j=1}^{n} \lambda_j Y_j\right)=\sum_{i=1}^{m} k_i \left(X_i,\ \sum_{j=1}^{n} \lambda_j Y_j\right)$$

$$=\sum_{i=1}^{m} k_i\left[\sum_{j=1}^{n} \lambda_j (X_i,\ Y_j)\right]$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{n} k_i \lambda_j (X_i,\ Y_j)$$

二、欧氏空间中向量的长度及两向量的夹角

由解析几何知道，R^3是一欧氏空间，其中对向量的长度及两向量夹角的定义，可以推广到一般的欧氏空间中来，即

定义2 由于欧氏空间V中任一向量X,总有$(X,X)\geqslant 0$。依此，称其算术根$\sqrt{(X,\ X)}$为向量X的长度，记为$|X|$，即

$$|X|=\sqrt{(X,\ X)}$$

当$|X|=1$时，又称X为单位向量。

向量的长度具有下述性质，即

定理1 设V是欧氏空间，则对V中任意向量X，Y及R中任意数k，有

(1) 齐次性：$|kX|=|k||X|$

(2) 非负性：$|X|\geqslant 0$，当且仅当$X=0$时，$|X|=0$

(3) Cauchy-Schwarz不等式：

$$(X,\ Y)^2\leqslant(X,\ X)(Y,\ Y)$$

或
$$|(X,\ Y)|\leqslant|X|\ |Y|$$

(4) 三角不等式：

$$|X+Y|\leqslant|X|+|Y| \qquad |X-Y|\geqslant|X|-|Y|$$

证明 性质(1)、(2)显然成立。下面只证明性质(3)，(4)。

对于性质(3)：当$Y=0$时，显然成立。今假设$Y\neq 0$，则对于任意$c\in R$，有

$$0\leqslant|Y-cY|^2=(X-cY,\ X-cY)$$
$$=(X,X)-2c(X,Y)+c^2(Y,\ Y)$$

由于$Y\neq 0$，故可取$c=(X,\ Y)/(Y,\ Y)$代入上式，得

$$0\leqslant(X,\ X)-\frac{(X,\ Y)^2}{(Y,\ Y)}$$

即
$$(X,\ Y)^2\leqslant(X,\ X)(Y,\ Y)$$

对于性质(4)，因

$$|X+Y|^2=(X+Y,\ X+Y)$$
$$=(X,\ X)+2(X,\ Y)+(Y,\ Y)$$

由Cauchy-Schwarz不等式，有

$$|X+Y|^2\leqslant|X|^2+2|X||Y|+|Y|^2=(|X|+|Y|)^2$$

即 $|X+Y|\leqslant|X|+|Y|$ 由此

$$|X|=|X-Y+Y|\leqslant|X-Y|+|Y|$$

即得 $|X-Y|\geqslant|X|-|Y|$ 证毕

还须指出：(3)及(4)中等号成立的充要条件是X，Y线性相关（读者自证）。

此外，将Cauchy-Schwarz不等式应用到例1和例2的欧氏空间中，可顺次得到常见的两个不等式：

$$(x_1y_1+x_2y_2+\cdots+x_ny_n)^2\leqslant(x_1^2\cdots+x_2^2+x_n^2)(y_1^2+y_2^2+\cdots+y_n^2)$$

及
$$\left(\int_a^b f(x)g(x)dx\right)^2\leqslant\int_a^b f^2(x)dx\cdot\int_a^b g^2(x)dx$$

有了定理1，我们就可以在欧氏空间中定义两个非零向量的夹角。

定义3 欧氏空间V中任意两个非零向量X，Y的夹角定义为

$$\varphi=\arccos\frac{(X,\ Y)}{|X||Y|}\quad 或\cos\varphi=\frac{(X,\ Y)}{|X||Y|}$$

若$(X,\ Y)=0$，则称X与Y正交（直交或垂直），记为$X\perp Y$，零向量与任意向量正交。

一般地，如向量α_1，α_2，…，α_m两两正交，则有

$$|\alpha_1+\alpha_2+\cdots+\alpha_m|^2=|\alpha_1|^2+|\alpha_2|^2+\cdots+|\alpha_m|^2$$

三、欧氏空间中内积与基底的关系

设e_1，e_2，…，e_n是n维欧氏空间V的一组基底，对于V中任意两向量

$$X=x_1e_1+x_2e_2+\cdots+x_ne_n\quad Y=y_1e_1+y_2e_2+\cdots+y_ne_n$$

由内积的性质可得

$$(X,\ Y)=(x_1e_1+x_2e_2+\cdots+x_ne_n,\ y_1e_1+y_2e_2+\cdots+y_ne_n)$$
$$=\sum_{i=1}^n\sum_{j=1}^n(e_i,\ e_j)x_iy_j$$

这就是内积与基底之间的关系。若令

$$a_{ij}=(e_i,\ e_j) \qquad (i,\ j=1,\ 2,\ \cdots,\ n)$$

再根据内积定义，显然有$a_{ij}=a_{ji}$，于是

$$(X,\ Y)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_iy_j$$

写成矩阵形式，则有

$$(X,\ Y)=XAY^{\mathrm{T}}$$

其中 $X=(x_1,\ x_2,\ \cdots,\ x_n) \qquad Y=(y_1,\ y_2,\ \cdots,\ y_n)$

$$A=[a_{ij}]=\begin{pmatrix}(e_1,\ e_1) & (e_1,\ e_2) & \cdots & (e_1,\ e_n)\\ (e_2,\ e_1) & (e_2,\ e_2) & \cdots & (e_2,\ e_n)\\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots\\ (e_n,\ e_1) & (e_n,\ e_2) & \cdots & (e_n,\ e_n)\end{pmatrix}$$

从上述得知，矩阵A是一个对称正定矩阵，称为V对于基底e_1，e_2，$\cdots$，e_n的**度量矩阵**。显然，知道了一组基底的度量矩阵，则任意两个向量的内积就可以通过坐标来计算。从而度量矩阵完全确定了内积。

四、正交矩阵与正交变换

从向量的正交概念，可以引出欧氏空间中的正交基底与标准正交基底，从而导出正交矩阵的概念。

定义4 欧氏空间V中的一组非零向量，如果它们两两正交，则称为正交向量组。

不难证明，正交向量组是线性无关的。因此，在n维欧氏空间中，两两正交的非零向量不超过n个。例如在平面上不存在三个两两正交的非零向量；在空间中不存在四个两两正交的非零向量。

定义5 若n维欧氏空间V的一组基底为正交向量组，则

称它们为V的一组正交基底。由单位向量组成的正交基底，则称为**标准正交基底**。

〔**例4**〕 设有定义在区间〔0，2π〕上的$2n+1$个连续函数：

1，$\cos x$，$\sin x$，$\cos 2x$，$\sin 2x$，…，$\cos nx$，$\sin nx$ 其线性组合

$$P(x)=a_0+a_1\cos x+b_1\sin x+\cdots+a_n\cos nx+b_n\sin nx$$

称为n次三角多项式。这种多项式的全体对通常的函数加法与数乘运算在实数域上构成一个$2n+1$维实线性空间R，又叫做n次三角多项式空间。其中的任意两个向量$P(x)$和$Q(x)$，若规定

$$(P(x),\ Q(x))=\int_0^{2\pi}P(x)Q(x)\mathrm{d}x$$

则可验证$(P(x),\ Q(x))$为内积，从而R为一欧氏空间。

由于 $$\int_0^{2\pi}\cos kx\cdot\cos lx\mathrm{d}x=0 \qquad (k\neq l)$$

$$\int_0^{2\pi}\sin kx\cdot\sin lx\mathrm{d}x=0 \qquad (k\neq l)$$

$$\int_0^{2\pi}\sin kx\cdot\cos lx\mathrm{d}x=0 \quad (k=1,\ 2,\ \cdots,\ n;\ l=0,\ 1,2,\ \cdots,\ n)$$

所以函数组

1，$\cos x$，$\sin x$，…，$\cos nx$，$\sin nx$

为R中的一组正交基底。再由积分

$$\int_0^{2\pi}\sin^2 kx\mathrm{d}x=\int_0^{2\pi}\cos^2 kx\mathrm{d}x=\pi \quad (k=1,\ 2,\ \cdots,\ n)$$

$$\int_0^{2\pi} dx = 2\pi$$

可得R中的一组标准正交基底：

$$\frac{1}{\sqrt{2\pi}},\ \frac{\cos x}{\sqrt{\pi}},\ \frac{\sin x}{\sqrt{\pi}},\ \cdots,\ \frac{\cos nx}{\sqrt{\pi}},\ \frac{\sin nx}{\sqrt{\pi}}$$

一般将欧氏空间中的一组基底化为正交基底和标准正交基底的方法，在《线性代数》中已讲过，这里不再重述。

此外，若e_1，e_2，…，e_n是欧氏空间V中的一组标准正交基底，由定义即有

$$(e_i, e_j)=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

就是说，一组基底为标准正交基底的充要条件是它的度量矩阵为单位矩阵。

设e_1，e_2，…，e_n与e'_1，e'_2，…，e'_n是欧氏空间V中的两组标准正交基底，它们的过渡矩阵为$A=[a_{ij}]$即有

$$\begin{pmatrix}e_1'\\ e_2'\\ \vdots\\ e_n'\end{pmatrix}=A^T\begin{pmatrix}e_1\\ e_2\\ \vdots\\ e_n\end{pmatrix}=\begin{pmatrix}a_{11}a_{21}\cdots a_{n1}\\ a_{12}a_{22}\cdots a_{n2}\\ \cdots\cdots\cdots\cdots\\ a_{1n}a_{2n}\cdots a_{nn}\end{pmatrix}\begin{pmatrix}e_1\\ e_2\\ \vdots\\ e_n\end{pmatrix}$$

则 $(e'_i, e'_j)=\left(\sum_{k=1}^{n} a_{ki}e_k,\ \sum_{k=1}^{n} a_{kj}e_k\right)\ (i,j=1,2,\cdots,n)$

由于 $(e_i', e_j')=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$

及

$$(e_i, e_j)=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

再利用内积性质，上式就成为

$$a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\begin{cases}1 & i=j\\0 & i\neq j\end{cases}$$

即

$$\begin{pmatrix}a_{11} & a_{21} & \cdots & a_{n1}\\a_{12} & a_{22} & \cdots & a_{n2}\\\cdots & \cdots & \cdots & \cdots\\a_{1n} & a_{2n} & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\a_{21} & a_{22} & \cdots & a_{2n}\\\cdots & \cdots & \cdots & \cdots\\a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}=\begin{pmatrix}1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$$

$$A^TA=I$$

或

$$A^{-1}=A^T$$

由此可引出正交矩阵的概念。

定义6 若n阶实矩阵A和它的转置矩阵A^T满足下列关系

$$A^TA=AA^T=I \quad 或 \quad A^{-1}=A^T$$

则称A为**正交矩阵**。

据此，上述结果表明：由标准正交基底到标准正交基底的过渡矩阵是正交矩阵。反之，若第一组基底是标准正交的，且过渡矩阵是正交矩阵，则第二组基底也是标准正交的。

此外，由定义6可以看出：矩阵A为正交矩阵的充要条件是A的列向量组为两两正交的单位向量组。即有

$$a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\begin{cases}1 & i=j\\0 & i\neq j\end{cases}$$

定义7 设$\mathscr{A}$是欧氏空间V中的线性变换，若对V中任意二向量X，Y恒有

$$(\mathscr{A}(X),\mathscr{A}(Y))=(X,Y)$$

（即保持向量的内积不变），则称$\mathscr{A}$为V中的一个正交变换。

〔**例5**〕 设$\mathscr{A}$是欧氏空间R^2中将向量绕原点旋转θ角的一个变换，则$\mathscr{A}$是R^2中的一个正交变换。

证明 在R^2中取直角坐标轴上的单位向量i，j为一组基底，则有

$$\begin{cases}\mathscr{A}(i)=\cos\theta i+\sin\theta j\\ \mathscr{A}(j)=-\sin\theta i+\cos\theta j\end{cases}$$

于是$\mathscr{A}$在基底i，j下的矩阵

$$A=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$

设X，$Y\in R^2$，记X，Y与$\mathscr{A}(X)$，$\mathscr{A}(Y)$在基底i，j下的坐标为

$$X=\begin{pmatrix}x_1\\ x_2\end{pmatrix},\quad Y=\begin{pmatrix}y_1\\ y_2\end{pmatrix},\quad \mathscr{A}(X)=\begin{pmatrix}x_1'\\ x_2'\end{pmatrix},\quad \mathscr{A}(Y)=\begin{pmatrix}y_1'\\ y_2'\end{pmatrix}$$

则有

$$\begin{pmatrix}x_1'\\ x_2'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\quad \begin{pmatrix}y_1'\\ y_2'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}$$

依此，有

$$(\mathscr{A}(X),\mathscr{A}(Y))=x_1'y_1'+x_2'y_2'=x_1y_1+x_2y_2=(X,Y)$$

所以$\mathscr{A}$是R^2中的一个正交变换。

正交变换有下述性质

定理2 设$\mathscr{A}$是欧氏空间V中的一个线性变换，则下面四个命题是等价的。

(1) $\mathscr{A}$是正交变换；

(2) $\mathscr{A}$保持向量的长度不变。即对任意$X\in V$，有

$|\mathscr{A}(X)|=|X|$；

(3) 若e_1，e_2，…，e_n是标准正交基底，则$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$也是标准正交基底；

(4) $\mathscr{A}$在任一组标准正交基底下的矩阵是正交矩阵。

证明 首先证明(1)与(2)等价。

如果$\mathscr{A}$是正交变换，则有

$$(\mathscr{A}(X),\ \mathscr{A}(X))=(X,X)$$

两边开方，即得

$$|\mathscr{A}(X)|=|X|$$

反过来，如果$\mathscr{A}$保持向量长度不变，则有

$$(\mathscr{A}(X),\mathscr{A}(X))=(X,X)$$

$$(\mathscr{A}(Y),\mathscr{A}(Y))=(Y,Y)$$

$$(\mathscr{A}(X+Y),\mathscr{A}(X+Y))=(X+Y,X+Y)$$

将最后这个等式两边展开，得

$$(\mathscr{A}(X),\mathscr{A}(X))+2(\mathscr{A}(X),\mathscr{A}(Y))+(\mathscr{A}(Y),\mathscr{A}(Y))$$
$$=(X,X)+2(X,Y)+(Y,Y)$$

利用前两个等式，即得

$$(\mathscr{A}(X),\mathscr{A}(Y))=(X,Y)$$

所以$\mathscr{A}$是正交变换。

再证(1)与(3)等价。

设e_1，e_2，…，e_n是一组标准正交基底，即

$$(e_i,e_j)=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}\quad (i,j=1,2,\cdots,n)$$

如果$\mathscr{A}$是正交变换，则有

$$(\mathscr{A}(e_i),\mathscr{A}(e_j))=(e_i,e_j)=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$
$$(i,j=1,2,\cdots,n)$$

所以$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$也是标准正交基底。

反过来，如果$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$是标准正交基底，则由

$$X=x_1e_1+x_2e_2+\cdots+x_ne_n$$

$$Y=y_1e_1+y_2e_2+\cdots+y_ne_n$$

和 $$\mathscr{A}(X)=x_1\mathscr{A}(e_1)+x_2\mathscr{A}(e_2)+\cdots+x_n\mathscr{A}(e_n)$$

$$\mathscr{A}(Y)=y_1\mathscr{A}(e_1)+y_2\mathscr{A}(e_2)+\cdots+y_n\mathscr{A}(e_n)$$

即得

$$(X,Y)=x_1y_1+x_2y_2+\cdots+x_ny_n=(\mathscr{A}(X),\mathscr{A}(Y))$$

因此$\mathscr{A}$是正交变换。

最后证明(3)与(4)等价。

设$\mathscr{A}$在标准正交基底e_1，e_2，…，e_n下的矩阵为A，即

$$\begin{pmatrix}\mathscr{A}(e_1)\\\mathscr{A}(e_2)\\\vdots\\\mathscr{A}(e_n)\end{pmatrix}=A^{\mathrm{T}}\begin{pmatrix}e_1\\e_2\\\vdots\\e_n\end{pmatrix}$$

若$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$是标准正交基底，则A可看作是由标准正交基底e_1，e_2，…，e_n到标准正交基底$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$的过渡矩阵，所以A是正交矩阵。

反过来，如果A是正交矩阵，则$\mathscr{A}(e_1)$，$\mathscr{A}(e_2)$，…，$\mathscr{A}(e_n)$就是标准正交基底。

这样，我们就证明了(1)，(2)，(3)，(4)的等价性。

〔例6〕 设R^2中的一个线性变换$\mathscr{A}$，在标准正交基底i，j(位于直角坐标轴上的单位向量)下的矩阵为

$$P=(I-2UU^{\mathrm{T}})J$$

其中 $U=\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$ $\qquad J=\begin{bmatrix}-1&0\\0&1\end{bmatrix}$

试证$\mathscr{A}$为正交变换。

证明 由于

$$P=\left(\begin{pmatrix}1&0\\0&1\end{pmatrix}-2\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}[\cos\theta\ \sin\theta]\right)\begin{pmatrix}-1&0\\0&1\end{pmatrix}$$

$$=\begin{pmatrix} -\cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}$$

则有 $\quad P^{T}P=I$

即P为正交矩阵，从而$\mathscr{A}$是一个正交变换。易见，它是将R^2中的向量绕原点旋转2θ角的变换。

五、酉空间

我们知道欧氏空间是实内积空间，即是引进了内积的实线性空间。现在我们在复线性空间中来引进内积，从而得出复内积空间，亦称酉空间。

定义8 设V是复数域C上的线性空间，若V中任意二向量α，β都按某一确定的法则对应于唯一确定的复数，记为(α,β)，它满足下列条件：

(1) $(\alpha,\beta)=\overline{(\beta,\alpha)}$

(2) $(k\alpha,\beta)=k(\alpha,\beta) \qquad (k\in C)$

(3) $(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$

(4) $(\alpha,\alpha)\geqslant 0$，当且仅当$\alpha=0$时$(\alpha,\alpha)=0$则称$(\alpha,\beta)$为向量$\alpha$与$\beta$的内积。定义了内积的复线性空间$V$，叫做复内积空间，亦称**酉空间**。酉空间和欧氏空间统称为内积空间。

注意：这里定义的内积与欧氏空间中定义的内积仅条件(1)有所不同。原因是这里的(α,β)一般是复数。但根据条件(1) (α,β)就是实数了。这样，条件(4)才有意义。据此，我们称

$$|\alpha|=\sqrt{(\alpha,\alpha)}$$

为向量α的长度。

一般在酉空间中不再定义两向量的夹角，但是当$(\alpha,\beta)=0$时，仍称向量α与β正交，并记为$\alpha\perp\beta$。

〔例7〕 若在n维复线性空间C^n中，对任意两向量$\alpha=(a_1, a_2, \cdots, a_n)$ $\beta=(b_1, b_2, \cdots, b_n)$ 规定

$$(\alpha,\beta)=a_1\overline{b_1}+a_2\overline{b_2}+\cdots+a_n\overline{b_n}$$

不难验证它满足定义8中的4个条件，即(α,β)是内积，从而C^n成为酉空间。

由酉空间的定义，可以推出下列性质

性质1 $(\alpha,\lambda\beta)=\overline{\lambda}(\alpha,\beta)$

性质2 $(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$

性质3 $(0,\alpha)=(\alpha,0)=0$

性质4 $\left(\sum\limits_{i=1}^{m}k_i\alpha_i,\ \sum\limits_{j=1}^{n}\lambda_j\beta_j\right)=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}k_i\overline{\lambda_j}(\alpha_i,\beta_j)$

性质5 Cauchy-Schwarz不等式

$$(\alpha,\beta)\overline{(\alpha,\beta)}\leqslant(\alpha,\alpha)(\beta,\beta)$$

证明 （性质1）

事实上 $(\alpha,\lambda\beta)=\overline{(\lambda\beta,\alpha)}=\overline{\lambda(\beta,\alpha)}=\overline{\lambda}(\alpha,\beta)$

（性质2）

事实上 $(\alpha,\beta+\gamma)=\overline{(\beta+\gamma,\alpha)}=\overline{(\beta,\alpha)}+\overline{(\gamma,\alpha)}$
$=(\alpha,\beta)+(\alpha,\gamma)$

（性质3） 由于

$$(0,\alpha)=(0\alpha,\alpha)=0(\alpha,\alpha)=0$$

故有 $(\alpha,0)=\overline{(0,\alpha)}=0$

（性质4） 事实上

$$\left(\sum_{i=1}^{m}k_i\alpha_i,\ \sum_{j=1}^{n}\lambda_j\beta_j\right)=\sum_{i=1}^{m}k_i\left(\alpha_i,\ \sum_{j=1}^{n}\lambda_j\beta_j\right)$$

$$=\sum_{i=1}^{m}k_i\sum_{j=1}^{n}\overline{\lambda_j}(\alpha_i,\ \beta_j)=\sum_{i=1}^{m}\sum_{j=1}^{n}k_i\overline{\lambda_j}(\alpha_i,\beta_j)$$

（性质5） 当$\beta=0$时，显然等号成立

现设$\beta\neq0$，对任意复数c，恒有

$$(\alpha-c\beta,\ \alpha-c\beta)\geqslant0$$

根据内积的性质1及性质2，有

$$(\alpha,\alpha)-\overline{c}(\alpha,\beta)-c(\beta,\alpha)+c\overline{c}(\beta,\beta)\geqslant0$$

取 $c=\dfrac{(\alpha,\beta)}{(\beta,\beta)}$代入上式，得

$$(\alpha,\alpha)-\frac{(\alpha,\beta)\overline{(\alpha,\beta)}}{(\beta,\beta)}\geqslant0$$

所以 $(\alpha,\beta)\overline{(\alpha,\beta)}\leqslant(\alpha,\alpha)(\beta,\beta)$ 证毕。

设$\alpha=(a_1,a_2,\cdots,a_n)$，$\beta=(b_1,b_2,\cdots,b_n)$为酉空间$C^n$中任意两向量，按例7中所规定的内积，应用Cauchy-Schwarz不等式，可得

$$\left(\sum_{i=1}^{n}a_i\overline{b_i}\right)\left(\sum_{i=1}^{n}\overline{a_i}b_i\right)\leqslant\left(\sum_{i=1}^{n}a_i\overline{a_i}\right)\left(\sum_{i=1}^{n}b_i\overline{b_i}\right)$$

六、酉空间中内积与基底的关系

与欧氏空间中的情形类似。设V是一个n维酉空间。在V中任取一组基底e_1，e_2，$\cdots$，e_n，对于V中任意两向量

$$\alpha=x_1e_1+x_2e_2+\cdots+x_ne_n$$
$$\beta=y_1e_1+y_2e_2+\cdots+y_ne_n$$

由内积的性质，有

$$(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\overline{y_j}(e_i,e_j)$$

这就是酉空间中内积与基底的关系。若令

$$a_{ij}=(e_i,e_j)\qquad(i,j=1,2,\cdots,n)$$

则显然有

$$a_{ij}=\overline{a_{ji}},\qquad 由此$$

$$(\alpha,\beta)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_i\overline{y_j}.$$

写成矩阵形式，则有

$$(\alpha,\beta)=XA\overline{Y}^{T}$$

其中　$X=[x_1,x_2,\cdots,x_n]$，$Y=[y_1,y_2,\cdots,y_n]$

$$A=\begin{pmatrix}(e_1,e_1) & (e_1,e_2)\cdots(e_1,e_n)\\ (e_2,e_1) & (e_2,e_2)\cdots(e_2,e_n)\\ \cdots\cdots & \cdots\cdots\cdots\cdots\\ (e_n,e_1) & (e_n,e_2)\cdots(e_n,e_n)\end{pmatrix}$$

矩阵A叫做V对于基底e_1，e_2，$\cdots$，e_n的度量矩阵。显然它满足$A=\overline{A}^T$，这里$\overline{A}=[\overline{a_{ij}}]$。

和实矩阵类似，若复矩阵A满足$A=\overline{A}^T$，则称A为**厄米特(Hermite)矩阵**。实的厄米特矩阵就是对称矩阵。

又若复矩阵A满足$\overline{A}^TA=A\overline{A}^T=I$或$A^{-1}=\overline{A}^T$，则称$A$为酉交矩阵(简称酉矩阵)，实的酉矩阵就是正交矩阵。

第三节　矩阵的Jordan标准形

在§1中曾经讲过，对于n阶方阵A与B，若存在n阶可逆方阵P，使得

$$B=P^{-1}AP$$

则称A与B相似。记为$A\sim B$。现在我们要来讨论：一切与已知矩阵A相似的矩阵中最简单的一种形式——Jordan标准形。为此，我们先介绍λ-矩阵。

一、λ-矩阵

以λ为变量的多项式，如

$$a(\lambda)=a_n\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$$

叫做λ的多项式。当$a_n=1$时，又称为**首一多项式**(Monic Polynomial)。以λ的多项式为元素的矩阵，如

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda)\cdots a_{1n}(\lambda) \\ a_{21}(\lambda) & a_{22}(\lambda)\cdots a_{2n}(\lambda) \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ a_{m1}(\lambda) & a_{m2}(\lambda)\cdots a_{mn}(\lambda) \end{pmatrix}$$

叫做多项式矩阵，又称为**λ-矩阵**。显然，λ-矩阵的任何子式计算出来，都是一个λ的多项式。

定义1 如果λ-矩阵$A(\lambda)$中有一个$r(\geqslant 1)$阶子式不恒为零，而所有的$r+1$阶子式(如果有的话)全部恒为零，则称$A(\lambda)$的秩为r，记作$\operatorname{rank}A(\lambda)=r$，零矩阵的秩为零。

定义2 设$A(\lambda)$是一个n阶λ-方阵，若存在n阶λ-方阵$B(\lambda)$，使得

$$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I$$

其中I是n阶单位矩阵，则称λ-矩阵$A(\lambda)$是可逆的，并称$B(\lambda)$是$A(\lambda)$的逆矩阵，记为$A^{-1}(\lambda)$。

容易证明：若$A(\lambda)$可逆，则其逆矩阵$A^{-1}(\lambda)$是唯一的。

在数字矩阵中，满秩矩阵是可逆的。但满秩的λ-矩阵，却不一定是可逆的。λ-矩阵可逆有如下的条件：

定理1 一个n阶λ-矩阵$A(\lambda)$可逆的充要条件是行列式$|A(\lambda)|$为一个非零常数。

证明 设$A(\lambda)$可逆，则有$B(\lambda)$使

$$A(\lambda)B(\lambda)=B(\lambda)A(\lambda)=I$$

两边取行列式，得

$$|A(\lambda)||B(\lambda)|=|I|=1$$

而$|A(\lambda)|$与$|B(\lambda)|$都是λ的多项式，上式表明其乘积为零次多项式1，故知它们也都是零次多项式，即$|A(\lambda)|$为非零常数。

反之，设$|A(\lambda)|=d$为非零常数，记$A(\lambda)$的伴随矩阵为$A^*(\lambda)$，则$\frac{1}{d}A^*(\lambda)$也是一个λ-矩阵，而 $A(\lambda)\frac{1}{d}A^*(\lambda)$

$$=\frac{1}{d}A^*(\lambda)A(\lambda)=I$$

所以$A(\lambda)$是可逆的，其逆矩阵$A^{-1}(\lambda)=\frac{1}{d}A^*(\lambda)$

此外，λ-矩阵有和数字矩阵类似的初等变换概念。即

定义3 下面三种变换

(1) 互换两行的位置；

(2) 把某一行乘以非零常数k；

(3) 把某一行乘以多项式$\varphi(\lambda)$加到另一行上去。

叫做λ-矩阵的**初等行变换**。把其中的"行"换为"列"即得λ-矩阵的**初等列变换**定义。初等行变换与初等列变换，统称为**初等变换**。

如果λ-矩阵$A(\lambda)$经过有限次初等变换化成$B(\lambda)$，就称$A(\lambda)$与$B(\lambda)$**等价**，记作$A(\lambda)\cong B(\lambda)$

可以验证等价关系满足

(1) 自反律：$A(\lambda)\cong A(\lambda)$；

(2) 对称律：若$A(\lambda)\cong B(\lambda)$，则$B(\lambda)\cong A(\lambda)$；

(3) 传递律：若$A(\lambda)\cong B(\lambda)$，$B(\lambda)\cong C(\lambda)$，则$A(\lambda)\cong C(\lambda)$

定理2 两个$m\times n$阶的λ-矩阵$A(\lambda)$，$B(\lambda)$等价的充要条

件是存在两个可逆矩阵$P(\lambda)$与$Q(\lambda)$使得

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$$

证明从略。

定理3 设A，B是两个n阶数字方阵，则$A\sim B$的充要条件是

$$\lambda I-A\cong\lambda I-B$$

证明 必要性是显然的。因为若$A\sim B$，则存在可逆矩阵P，使得$B=P^{-1}AP$，于是有

$$\lambda I-B=\lambda I-P^{-1}AP=P^{-1}(\lambda I-A)P$$

由定理2 即知

$$\lambda I-B\cong\lambda I-A$$

（充分性的证明从略）

这个定理的作用，在于沟通了两个数字矩阵的相似关系与其所对应的特征矩阵的等价关系，使我们可以通过等价关系来研究相似关系。

定理4 任意一个秩为r，且是非零的$m\times n$阶λ-矩阵$A(\lambda)$都等价于一个对角形矩阵

$$\begin{pmatrix} d_1(\lambda) & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & d_2(\lambda) & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & d_r(\lambda) & & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix}$$

其中$d_i(\lambda)$都是首一多项式，且$d_{i-1}(\lambda)$能整除$d_i(\lambda)$，记作$d_{i-1}(\lambda)|d_i(\lambda)$ $(i=2, 3, \cdots, r)$，这种对角形矩阵叫做$A(\lambda)$的**Smith标准形**。

本定理我们也略去证明，下面举例说明如何用初等变换将一个非零λ-矩阵化为它的Smith标准形。同时为了表明每一步中所用的为何种初等变换，我们采用如下记号：

(1) $[i,j]$——表示互换i，j两行(列)的位置；

(2) $[i(k)]$——表示将第i行(列)乘以非零常数k；

(3) $[i+j(\varphi)]$——表示将第j行(列)乘以多项式$\varphi(\lambda)$后，再加到第i行(列)上去。

而且，这些记号若写在"$\longrightarrow$"的上面，表示行变换；写在其下面，则表示列变换。

〔例1〕 用初等变换化λ-矩阵

$$A(\lambda)=\begin{pmatrix}1-\lambda & \lambda^2 & \lambda\\ \lambda & \lambda & -\lambda\\ 1+\lambda^2 & \lambda^2 & -\lambda^2\end{pmatrix}$$

为Smith标准形。

解 $$A(\lambda)\xrightarrow[{[1+3]}]{}\begin{pmatrix}1 & \lambda^2 & \lambda\\ 0 & \lambda & -\lambda\\ 1 & \lambda^2 & -\lambda^2\end{pmatrix}\xrightarrow{[3-1]}\begin{pmatrix}1 & \lambda^2 & \lambda\\ 0 & \lambda & -\lambda\\ 0 & 0 & -\lambda^2-\lambda\end{pmatrix}$$

$$\xrightarrow[{[2-1(\lambda^2)]}]{}\begin{pmatrix}1 & 0 & \lambda\\ 0 & \lambda & -\lambda\\ 0 & 0 & -\lambda^2-\lambda\end{pmatrix}\xrightarrow[{[3-1(\lambda)]}]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda & -\lambda\\ 0 & 0 & -\lambda^2-\lambda\end{pmatrix}$$

$$\xrightarrow[{[3+2]}]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & -\lambda^2-\lambda\end{pmatrix}\xrightarrow{[3(-1)]}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda(\lambda+1)\end{pmatrix}$$

二、行列式因子、不变因子和初等因子

定义4 设$\operatorname{rank}A(\lambda)=r(\geqslant 1)$，则对任一正整数$k$、只要$1\leqslant k\leqslant r$，$A(\lambda)$中必有非零的$k$阶子式。现将$A(\lambda)$中全部$k$阶子式的最高公因式(为首一多项式的)，记为$D_k(\lambda)$，称之为

$A(\lambda)$的k阶**行列式因子**。

显然，当$\operatorname{rank}A(\lambda)=r\geqslant 1$时，$A(\lambda)$共有$r$个行列式因子$D_1(\lambda)$，$D_2(\lambda)$，…，$D_r(\lambda)$。注意到$A(\lambda)$的每个$k$阶子式都可以按一行展开为$k$个$k-1$阶子式的组合，因此有

$$D_{k-1}(\lambda)\mid D_k(\lambda)\qquad (k=2,3,\cdots,r)$$

由此又得到r个首一多项式

$$D_1(\lambda),\quad \frac{D_2(\lambda)}{D_1(\lambda)},\ \frac{D_3(\lambda)}{D_2(\lambda)},\ \cdots,\ \frac{D_r(\lambda)}{D_{r-1}(\lambda)}$$

称为λ-矩阵$A(\lambda)$的**不变因子**

定理5 设$A(\lambda)$，$B(\lambda)$是等价的$m\times n$阶λ-矩阵，则$A(\lambda)$与$B(\lambda)$有相同的秩，且有相同的行列式因子。

证明 只需证明经过一次初等行变换把$A(\lambda)$化成$B(\lambda)$后，其秩与行列式因子不变即可。

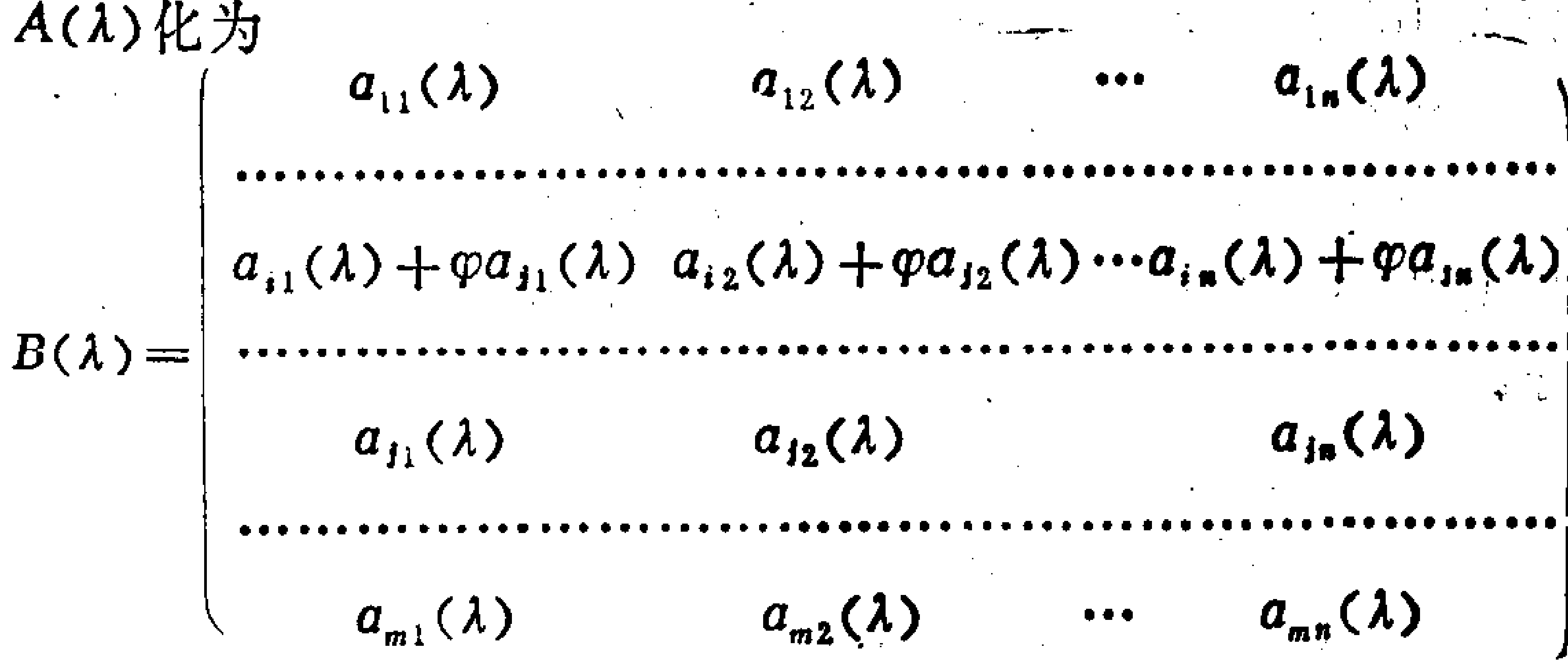
用$D_k^A(\lambda)$与$D_k^B(\lambda)$依次表示$A(\lambda)$与$B(\lambda)$的k阶行列式因子；用$m(\lambda)$与$n(\lambda)$依次表示在$A(\lambda)$与$B(\lambda)$中处于相同位置的任一k阶子式。

1° 假定经过的是第(1)种初等变换$[i,j]$，而将$A(\lambda)$化成$B(\lambda)$。那么，当$m(\lambda)$不含$A(\lambda)$的第i，j行时，有$n(\lambda)=m(\lambda)$；当$m(\lambda)$同时含有$A(\lambda)$的第i，j两行时，有$n(\lambda)=-m(\lambda)$；而当$m(\lambda)$只含有$A(\lambda)$的第i,j两行中之任一行时，则有$n(\lambda)=\pm m_1(\lambda)$，其中$m_1(\lambda)$是$A(\lambda)$中的另一个$k$阶子式。无论是哪一种，都有$D_k^A(\lambda)\mid n(\lambda)$，因而有$D_k^A(\lambda)\mid D_k^B(\lambda)$。

2° 假定经过的是第(2)种初等行变换$[i(k)]$，而将$A(\lambda)$化成$B(\lambda)$，则当$m(\lambda)$不含$A(\lambda)$的第i行时，显然有$n(\lambda)=m(\lambda)$；而当$m(\lambda)$含有$A(\lambda)$的第i行时，则必有$n(\lambda)=km(\lambda)$。因此总有$D_k^A(\lambda)\mid n(\lambda)$，所以有$D_k^A(\lambda)\mid D_k^B(\lambda)$。

3° 假定经过的是第(3)种初等行变换$[i+j(\varphi)]$，而将

$A(\lambda)$化为



$$B(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{i1}(\lambda)+\varphi a_{j1}(\lambda) & a_{i2}(\lambda)+\varphi a_{j2}(\lambda) & \cdots & a_{in}(\lambda)+\varphi a_{jn}(\lambda) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{j1}(\lambda) & a_{j2}(\lambda) & & a_{jn}(\lambda) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$$

则当$m(\lambda)$不含$A(\lambda)$的第i行时，有$n(\lambda)=m(\lambda)$；当$m(\lambda)$同时含有$A(\lambda)$的第i，j两行时，也有$n(\lambda)=m(\lambda)$；当$m(\lambda)$含$A(\lambda)$的第i行，但不含$A(\lambda)$的第j行时，则有$n(\lambda)=m(\lambda)\pm\varphi m_1(\lambda)$，这里$m_1(\lambda)$是$A(\lambda)$的另一个$k$阶子式。无论是哪一种，都有$D_k^A((\lambda)/n(\lambda)$，从而有$D_k^A(\lambda)/D_k^B(\lambda)$。

可见在$A(\lambda)\cong B(\lambda)$时，必有$D_k^A(\lambda)/D_k^B(\lambda)$。由对称律，在$A(\lambda)\cong B(\lambda)$时，亦有$B(\lambda)\cong A(\lambda)$，故亦有$D_k^B(\lambda)/D_k^A(\lambda)$。再由$D_k^A(\lambda)$与$D_k^B(\lambda)$都是首一多项式，所以必有$D_k^A(\lambda)=D_k^B(\lambda)$，即$A(\lambda)$与$B(\lambda)$有相同的行列式因子。

又当$A(\lambda)$的一切i阶子式均为零时，则有$D_i^A(\lambda)=0$，从而有$D_i^B(\lambda)=0$，于是$B(\lambda)$的一切i阶子式也均为零。反之亦然。这就是说，有$\mathrm{rank}A(\lambda)=\mathrm{rank}B(\lambda)$。定理证毕。

推论 设$A(\lambda)$的Smith标准形为

$$B(\lambda)=\begin{pmatrix} d_1(\lambda) & & & & & \\ & d_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & d_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

则其中$d_1(\lambda)$，$d_2(\lambda)$，…，$d_r(\lambda)$就是$A(\lambda)$的不变因子。

证明 由于$A(\lambda)\cong B(\lambda)$，则$A(\lambda)$与$B(\lambda)$有相同的秩和相同的各阶行列式因子。因此，$A(\lambda)$的秩，就是标准形$B(\lambda)$中主对角线上非零元素的个数$r$；$A(\lambda)$的各阶行列式因子，就是

$$D_1(\lambda)=d_1(\lambda)$$

$$D_2(\lambda)=d_1(\lambda)d_2(\lambda)$$

$$\cdots\cdots$$

$$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)$$

从而有 $d_1(\lambda)=D_1(\lambda)$

$$d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)}$$

$$\cdots\cdots$$

$$d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}$$

可见，$d_1(\lambda)$，$d_2(\lambda)$，…，$d_r(\lambda)$就是$A(\lambda)$的不变因子。因此也就说明了λ-矩阵的smith标准形是唯一的。

定理6 设$A(\lambda)$与$B(\lambda)$是$m\times n$阶的λ-矩阵，则下面命题是等价的。

(1) $A(\lambda)\cong B(\lambda)$；

(2) $A(\lambda)$，$B(\lambda)$有相同的行列式因子；

(3) $A(\lambda)$，$B(\lambda)$有相同的不变因子；

(4) $A(\lambda)$，$B(\lambda)$有相同的smith标准形；

(5) 存在可逆矩阵$P(\lambda)$与$Q(\lambda)$使

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda)$$

证明 用循环证法：

$$(1)\Longrightarrow(2)\Longrightarrow(3)\Longrightarrow(4)\Longrightarrow(1)\text{和}$$

(1)⟹(5)⟹(1)

(1)⟹(5)⟹(1)：见定理2

(1)⟹(2)：见定理5

(2)⟹(3)：因$A(\lambda)$，$B(\lambda)$有相同的行列与因子

$D_k^A(\lambda)=D_k^B(\lambda)\qquad(k=1,2,\cdots,r)$

所以有

$$d_i^A(\lambda)=\frac{D_i^A(\lambda)}{D_{i-1}^A(\lambda)}=\frac{D_i^B\lambda}{D_{i-1}^B(\lambda)}=d_i^B(\lambda),$$

$$(i=1,\ 2,\cdots,\ r)$$

(3)⇒(4)与(4)⇒(1)都是显然的

此定理表明，一个λ-矩阵经过初等变换后，不仅行列式因子不变，而且不变因子也不变。不变因子的得名，即由于此。

为了得出Jordan标准形，我们还需要进一步讨论λ-矩阵等价的充要条件。为此，再引进初等因子的概念。(下面在复数域上进行讨论)。

定义5 设rank$A(r)=r$，$A(\lambda)$的不变因子为$d_1(\lambda)$，$d_2(\lambda)$，…，$d_r(\lambda)$；$d_{i-1}(\lambda)|d_i(\lambda)$，$(i=2,3,\cdots,r)$。把$d_i(\lambda)$在复数域上分解成一次因式幂的乘积。

$$d_1(\lambda)=(\lambda-\lambda_1)^{k_{11}}(\lambda-\lambda_2)^{k_{12}}\cdots(\lambda-\lambda_j)^{k_{1j}}\cdots(\lambda-\lambda_t)^{k_{1t}}$$

……………………………………………………………

$$d_i(\lambda)=(\lambda-\lambda_1)^{k_{i1}}(\lambda-\lambda_2)^{k_{i2}}\cdots(\lambda-\lambda_j)^{k_{ij}}\cdots(\lambda-\lambda_t)^{k_{it}}$$

……………………………………………………………

$$d_r(\lambda)=(\lambda-\lambda_1)^{k_{r1}}(\lambda-\lambda_2)^{k_{r2}}\cdots(\lambda-\lambda_j)^{k_{rj}}\cdots(\lambda-\lambda_t)^{k_{rt}}$$

其中$k_{ij}\geqslant0$($i=1,\ 2,\ \cdots,\ r$；$j=1,\ 2,\ \cdots,\ t$)，称其中$k_{ij}>0$的一切$(\lambda-\lambda_j)^{k_{ij}}$为$A(\lambda)$的**初等因子**。

应当注意，$A(\lambda)$的初等因子可能有重复的，比如$(\lambda-\lambda_j)^m$可能既是$d_r(\lambda)$的一次因式幂，又是$d_{r-1}(\lambda)$的一次

因式幂。此时，$(\lambda-\lambda_j)^m$至少要算$A(\lambda)$的两个初等因子。通常将

$$(\lambda-\lambda_j)^{k_{1j}},\ \cdots,\ (\lambda-\lambda_j)^{k_{ij}},\ \cdots,\ (\lambda-\lambda_j)^{k_{rj}}$$

称为与$(\lambda-\lambda_j)$**相当的**初等因子。

定理7 两个$m\times n$阶的λ-矩阵$A(\lambda)$，$B(\lambda)$等价的充要条件是它们有相同的秩和相同的初等因子。

证明 必要性由定理5，定理6及定义5即得。今证充分性：不失一般性，我们用一个具体例子来说明：

假定$A(\lambda)$与$B(\lambda)$其秩均为4，且有相同的初等因子

$$\lambda,\ \lambda^2,\ \lambda^3,\ \lambda-1,\ \lambda-1,\ \lambda+1,\ (\lambda+1)^2$$

由于秩都是4，故它们都有4个不变因子$d_i^A(\lambda)$与$d_i^B(\lambda)$ $(i=1,2,3,4)$。现将初等因子中分别与λ，$\lambda-1$，$\lambda+1$相当的，各自按降幂排列，当不足4个时用1补成4个：

$$\begin{array}{llll}\lambda^3, & \lambda^2, & \lambda, & 1\\ \lambda-1, & \lambda-1, & 1, & 1\\ (\lambda+1)^2, & \lambda+1, & 1, & 1\end{array}$$

由于不变因子$d_{i-1}^A(\lambda)\mid d_i^A(\lambda)$，$d_{i-1}^B(\lambda)\mid d_i^B(\lambda)$，$(i=2,3,4)$可知$d_4^A(\lambda)$与$d_4^B(\lambda)$都等于上式第一列中各初等因子的乘积

$$d_4^A(\lambda)=d_4^B(\lambda)=\lambda^3(\lambda-1)(\lambda+1)^2$$

$d_3^A(\lambda)$与$d_3^B(\lambda)$均等于上式第二列中各初等因子的乘积

$$d_3^A(\lambda)=d_3^B(\lambda)=\lambda^2(\lambda-1)(\lambda+1)$$

类此有

$$d_2^A(\lambda)=d_2^B(\lambda)=\lambda$$

$$d_1^A(\lambda)=d_1^B(\lambda)=1$$

所以 $d_i^A(\lambda)=d_i^B(\lambda)\quad (i=1,2,3,4)$

由定理6即知 $A(\lambda)\cong B(\lambda)$

对于一个λ-矩阵，求出其初等因子，是化此λ-矩阵为Jor-

dan标准形的重要步骤。因此，我们必须掌握求初等因子的方法。

按定义，初等因子可以通过行列式因子和不变因子来求。

〔例2〕 求n阶λ-矩阵

$$A(\lambda)=\begin{pmatrix} \lambda-a & b_1 & & \\ & \lambda-a & \ddots & \\ & & \ddots & b_{n-1} \\ & & & \lambda-a \end{pmatrix}$$

的初等因子。这里b_1，b_2，…，b_{n-1}都是非零常数。

解 显然$D_n(\lambda)=(\lambda-a)^n$，又在$A(\lambda)$中划去第1列和第$n$行，得到其右上角的$n-1$阶子式，其值为非零常数$b_1b_2\cdots b_{n-1}$，所以$D_{n-1}(\lambda)=1$。再由$D_{i-1}(\lambda)|D_i(\lambda)$，可知$D_{n-2}(\lambda)=\cdots=D_1(\lambda)=1$因此

$$d_1(\lambda)=\cdots=d_{n-1}(\lambda)=1,\ d_n(\lambda)=(\lambda-a)^n$$

可见，$A(\lambda)$的初等因子只有一个$(\lambda-a)^n$

同样，形如

$$\begin{pmatrix} \lambda-a & & & \\ b_1 & \lambda-a & & \\ & \ddots & \ddots & \\ & & b_{n-1} & \lambda-a \end{pmatrix}$$

$(b_i\neq 0)$的λ-矩阵，也只有一个初等因子$(\lambda-a)^n$

此外，初等因子也可以通过对角形(不必需要Smith标准形)或分块对角形来求。这就是下面的两个定理(证明从略)。

定理8 若λ-矩阵

$$A(\lambda)\cong\begin{pmatrix} f_1(\lambda) & & & & & \\ & f_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & f_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix}$$

则其中λ多项式$f_1(\lambda)$，$f_2(\lambda)$，$\cdots$，$f_r(\lambda)$的所有一次因式幂，就是$A(\lambda)$的全部初等因子。

〔**例3**〕 求矩阵

$$\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$$

的特征矩阵的初等因子

解 $\lambda I-A=\begin{pmatrix} \lambda-4 & -6 & 0 \\ 3 & \lambda+5 & 0 \\ 3 & 6 & \lambda-1 \end{pmatrix}\xrightarrow{[3-2]}\begin{pmatrix} \lambda-4 & 6 & 0 \\ 3 & \lambda+5 & 0 \\ 0 & -\lambda+1 & \lambda-1 \end{pmatrix}$

$$\xrightarrow[{[2+3]}]{}\begin{pmatrix} \lambda-4 & -6 & 0 \\ 3 & \lambda+5 & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix}\xrightarrow{[1(3)-2(\lambda-4)]}$$

$$\begin{pmatrix} 0 & -\lambda^2-\lambda+2 & 0 \\ 3 & \lambda+5 & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix}\xrightarrow[2-1\left(\frac{\lambda+5}{3}\right)]{1(-1)}\begin{pmatrix} 0 & \lambda^2+\lambda-2 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix}\xrightarrow{[1,\ 2]}$$

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & (\lambda+2)(\lambda-1) & 0 \\ 0 & 0 & \lambda-1 \end{pmatrix}$$

根据定理8，$\lambda I-A$的初等因子为$\lambda+2$，$\lambda-1$，$\lambda-1$。

定理9 若λ-矩阵$A(\lambda)$呈分块对角形

$$A(\lambda)=\begin{pmatrix} A_1(\lambda) & & & \\ & A_2(\lambda) & & \\ & & \ddots & \\ & & & A_r(\lambda)\end{pmatrix}$$

则其中每个λ-矩阵$A_1(\lambda)$，$A_2(\lambda)$，…，$A_r(\lambda)$的初等因子，就是$A(\lambda)$的全部初等因式。

〔**例4**〕 求λ-矩阵

$$A(\lambda)=\begin{pmatrix} \lambda^2 & \lambda & & \\ 0 & 1 & & \\ & & 0 & \lambda \\ & & \lambda-2 & 1\end{pmatrix}$$

的初等因子。

解 因为$A(\lambda)$是分块对角形，其中

$$A_1(\lambda)=\begin{pmatrix} \lambda^2 & \lambda \\ 0 & 1\end{pmatrix}\xrightarrow{[1-2(\lambda)]}\begin{pmatrix} \lambda^2 & 0 \\ 0 & 1\end{pmatrix}$$

故其初等因子为λ^2，而

$$A_2(\lambda)=\begin{pmatrix} 0 & \lambda \\ \lambda-2 & 1\end{pmatrix}\xrightarrow[{[1-2(\lambda-2)]}]{}\begin{pmatrix} -\lambda(\lambda-2) & \lambda \\ 0 & 1\end{pmatrix}$$

$$\xrightarrow{[1-2(\lambda)]}\begin{pmatrix} -\lambda(\lambda-2) & 0 \\ 0 & 1\end{pmatrix}\xrightarrow{[1(-1)]}$$

$$\begin{pmatrix} \lambda(\lambda-2) & 0 \\ 0 & 1\end{pmatrix}$$

可见，其初等因子为λ，$\lambda-2$，所以$A(\lambda)$的全部初等因子是λ^2，λ，$\lambda-2$。

三、Jordan标准形

现在我们就来介绍与一个n阶矩阵A相似的最简单的一种形式——Jordan标准形。

定义6 设有n阶矩阵A，其特征矩阵$\lambda I-A$的初等因子是$(\lambda-\lambda_1)^{n_1}$，$(\lambda-\lambda_2)^{n_2}$，…，$(\lambda-\lambda_t)^{n_t}$，$(n_1+n_2+\cdots+n_t=n)$。称$n_i$阶矩阵

$$J_i=\begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}$$

为属于初等因子$(\lambda-\lambda_i)^{n_i}$的Jordan块$(i=1,2,\cdots,t)$由这些Jordan块构成的$n$阶分块对角矩阵

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix}$$

就叫做矩阵A的**Jordan标准形**。

定理10 在定义6的条件下，有$A\sim J$。

证明 根据定理3，只要证明$\lambda I-A\cong\lambda I-J$即可。由于特征矩阵总是满秩的，故有

$$\text{rank}[\lambda I-A]=\text{rank}[\lambda I-J]=n$$

又由例2得知

$$\lambda I_i-J_i=\begin{pmatrix} \lambda-\lambda_i & -1 & & & \\ & \lambda-\lambda_i & -1 & & \\ & & \ddots & \ddots & \\ & & & & -1 \\ & & & & \lambda-\lambda_i \end{pmatrix}_{n_i\times n_i}$$

（I_i为n_i阶单位矩阵）有唯一个初等因子$(\lambda-\lambda_i)^{n_i}$

依此，由定理9则n阶分块矩阵

$$\lambda I-J=\begin{pmatrix}\lambda I_1-J_1 & & & \\ & \lambda I_2-J_2 & & \\ & & \ddots & \\ & & & \lambda I_t-J_t\end{pmatrix}$$

的全部初等因子就是$(\lambda-\lambda_1)^{n_1}$，$(\lambda-\lambda_2)^{n_2}$，…，$(\lambda-\lambda_t)^{n_t}$。正好与$\lambda I-A$的初等因子相同，据定理7即知

$$\lambda I-A\cong\lambda I-J$$

从而有　$A\sim J$　　证毕

由于相似矩阵有相同的特征多项式，故有

$$|\lambda I-A|=|\lambda I-J|=D_n(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)$$
$$=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_t)^{n_t}$$

可见λ_1，λ_2，…，λ_t是矩阵A的特征值。但要注意：

当$i\neq j$时，可能有$\lambda_i=\lambda_j$。

再由于λ-矩阵的初等因子，不因初等变换而改变，而由λ-矩阵自身唯一决定。所以，属于特征矩阵$\lambda I-A$的每一个初等因子的Jordan块都是唯一的。从而，若不计J_i在矩阵A的Jordan标准形J中的顺序，则J也就是唯一的了。

〔例5〕 求矩阵

$$A=\begin{pmatrix}0 & 3 & 3\\ -1 & 8 & 6\\ 2 & -14 & -10\end{pmatrix}$$

的Jordan标准形J。

解
$$\lambda I-A=\begin{pmatrix}\lambda & -3 & -3\\ 1 & \lambda-8 & -6\\ -2 & 14 & \lambda+10\end{pmatrix}$$

$$\cong\begin{pmatrix}1 & & \\ & 1 & \\ & & \lambda(\lambda+1)^2\end{pmatrix}$$

其初等因子为λ，$(\lambda+1)^2$。故

$$A\sim J=\begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}$$

〔**例6**〕 求矩阵

$$A=\begin{pmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{pmatrix}$$

的Jordan标准形J。

解 由例3知$\lambda I-A$的初等因子为$\lambda+2$，$\lambda-1$，$\lambda-1$。所以

$$A\sim J=\begin{pmatrix}-2 & & \\ & 1 & \\ & & 1\end{pmatrix}$$

即A与对角形矩阵相似。

不难看出，一般当且仅当$\lambda I-A$的初等因子全是一次式时，A的 Jordan 标准形就是对角形(即是A相似于对角形矩阵)。

四、把A变成J的相似变换矩阵P

以上给出了将一个矩阵A化为Jordan标准形J的方法。由

于$A\sim J$，故必存在满秩的相似变换矩阵P，使得

$$P^{-1}AP=J$$

一般如只需求J时，就无须求P了。但是在一些具体问题中，如使用Jordan标准形求解微分方程时，就必须求出P来。

下面我们来介绍求矩阵P的方法，简单说来，就是将上式改写成

$$AP=PJ$$

令 $$P=[X_1,\ X_2,\ \cdots,\ X_n]$$

得 $$A[X_1,\ X_2,\ \cdots,\ X_n]=[X_1,\ X_2,\ \cdots,\ X_n]J$$

由此等式，再根据具体的A和J求出一组X_1，X_2，$\cdots$，X_n来作为P的列向量，就得到矩阵P。

例如：

$$J=\begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{pmatrix}$$

系由一个Jordan块组成，则有

$$A[X_1,\ X_2,\ \cdots,\ X_n]=[X_1,\ X_2,\ \cdots,\ X_n]\begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & \lambda \end{pmatrix}$$

由矩阵乘法规则，有

$[AX_1,\ AX_2,\ \cdots,\ AX_n]=[\lambda X_1,\ X_1+\lambda X_2, X_2+\lambda X_3,\ \cdots,\ X_{n-1}+\lambda X_n]$比较两端，得

$$
\begin{aligned}
AX_1&=\lambda X_1 & \text{或}\quad (\lambda I-A)X_1&=0\\
AX_2&=X_1+\lambda X_2 & (\lambda I-A)X_2&=-X_1\\
AX_3&=X_2+\lambda X_3 & (\lambda I-A)X_3&=-X_2\\
&\cdots\cdots & &\cdots\cdots\\
AX_n&=X_{n-1}+\lambda X_n & (\lambda I-A)X_n&=-X_{n-1}
\end{aligned}
$$

由于J中的λ是矩阵A的特征值，故由第一式知X_1就是A对于特征值λ的特征向量。求出X_1后，代入第二式，得一非齐次线性方程组，解出X_2再代入第三式求解X_3。如此下去，直至解出X_n来。由于矩阵P是存在的，故这些方程组一定都有解，而且不是唯一的。我们只要在其中选出一组线性无关的解来组成矩阵P就可以了，可见矩阵P也不是唯一的。

当J是由多个 Jordan 块组成时，求P的方法完全与此类似。

〔例7〕 求矩阵

$$
A=\begin{pmatrix} -1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2 \end{pmatrix}
$$

的Jordan标准形J，并求使$P^{-1}AP=J$的满秩矩阵P。

解

$$
\lambda I-A=\begin{pmatrix} \lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2 \end{pmatrix}
$$

其行列式因子 $D_3(\lambda)=|\lambda I-A|=\begin{vmatrix} \lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2 \end{vmatrix}$

按最后一列展开，即得

$$
D_3(\lambda)=(\lambda-2)(\lambda-1)^2
$$

又左下角的二阶子式 $\begin{vmatrix} 4 & \lambda-3 \\ -1 & 0 \end{vmatrix}=\lambda-3$

而$D_2(\lambda)\,|D_3(\lambda)$，故只有$D_2(\lambda)=1$，从而$D_1(\lambda)=1$。于是，不变因子为

$d_1(\lambda)=1$，$d_2(\lambda)=1$，$d_3(\lambda)=(\lambda-2)(\lambda-1)^2$，可见$\lambda I-A$的初等因子是$(\lambda-2)$，$(\lambda-1)^2$。所以

$$A\sim J=\begin{pmatrix} 2 & & \\ & 1 & 1 \\ & & 1 \end{pmatrix}$$

再来求P，令

$$P=[X_1,\ X_2,\ X_3]\quad 其中X_i=\begin{pmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \end{pmatrix}(i=1,\ 2,\ 3)$$

代入 $AP=PJ$

得 $$A[X_1,\ X_2,\ X_3]=[X_1,\ X_2,\ X_3]\begin{pmatrix} 2 & & \\ & 1 & 1 \\ & & 1 \end{pmatrix}$$

比较两端，得

$$\begin{aligned} &AX_1=2X_1 && (2I-A)X_1=0 \\ &AX_2=X_2 \quad 或 && (I-A)X_2=0 \\ &AX_3=X_2+X_3 && (I-A)X_3=-X_2 \end{aligned}$$

由第一式，得齐次线性方程组：$\begin{cases} 3x_{11}-x_{12}+0=0 \\ 4x_{11}-x_{12}+0=0 \\ -x_{11}+0+0=0 \end{cases}$

由此解得A对于特征值2的特征向量

$$X_1=\begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

又由第二式，得齐次线性方程组：

$$\begin{cases} 2x_{21}-x_{22} \quad\quad =0 \\ 4x_{21}-2x_{22} \quad\quad =0 \\ -x_{21} \quad\quad -x_{23}=0 \end{cases}$$

解得A对于特征值1的特征向量

$$X_2=\begin{pmatrix} x_{21} \\ x_{22} \\ x_{23} \end{pmatrix}=k\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

其中k为非零常数，其取值有赖于将X_2代入第三式后，使其方程有解来确定。据此，由第三式得非齐次线性方程组：

$$\begin{cases} 2x_{31}-x_{32} \quad\quad =-k \\ 4x_{31}-2x_{32} \quad\quad =-2k \\ -x_{31} \quad\quad -x_{33}=k \end{cases}$$

欲其有解，必须其系数矩阵与其增广矩阵，即

$$\begin{pmatrix} 2 & -1 & 0 \\ 4 & -2 & 0 \\ -1 & 0 & -1 \end{pmatrix} 与 \begin{pmatrix} 2 & -1 & 0 & -k \\ 4 & -2 & 0 & -2k \\ -1 & 0 & -1 & k \end{pmatrix}$$

的秩相等。不难看出，在这里k取任何非零常数均可。如取$k=1$，则有

$$X_2=\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \quad 和 \quad \begin{cases} 2x_{31}-x_{32} \quad\quad =-1 \\ 4x_{31}-2x_{32} \quad\quad =-2 \\ -x_{31} \quad\quad -x_{33}=1 \end{cases}$$

从此方程组解得

$$X_3=\begin{pmatrix} x_{31} \\ x_{32} \\ x_{33} \end{pmatrix}=k_1\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}+\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

如取$k_1=-1$，得

$$X_3=\begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}$$

于是 $P=[X_1,\ X_2,\ X_3]=\begin{pmatrix} 0 & 1 & -1 \\ 0 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}$

我们看到，其中的X_1与X_2分别是A对于特征值2和1的特征向量，而X_3一般则称为A对于特征值1的**广义特征向量**。

习 题 一

1. 验证下列集合对于所指的线性运算，是否构成实数域上的线性空间。

(1) 次数低于n的实系数多项式的全体，对于多项式的加法和数乘法。

(2) 平面上不平行于某一向量的全体向量组成的集合，对于向量的加法和数乘法。

2. 判别下列所定义的变换，哪些是线性的，哪些不是。

(1) 在三维线性空间中

$\mathscr{A}(x_1,\ x_2,\ x_3)=(x_1^2,\ x_2+x_3,\ x_3^2)$

$\mathscr{A}(x_1,\ x_2,\ x_3)=(2x_1-x_2,\ x_2+x_3,\ x_1)$

(2) 在$P[x]$（多项式的全体）中

$\mathscr{A}f(x)=f(x+1)$

$\mathscr{A}f(x)=f(x_0)$ 其中$x_0\in F$，x_0是一个定数。

3. 设V是n维实线性空间，$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n$ 是V中的一组基，对V中任意二向量

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$$

$$\beta=y_1\alpha_1+y_2\alpha_2+\cdots+y_n\alpha_n$$

规定　　$(\alpha,\ \beta)=x_1y_1+2x_2y_2+\cdots nx_ny_n$

证明：$(\alpha,\ \beta)$ 是α与β的内积，从而V对这个内积构成欧氏空间。

4. 设$A=[a_{ij}]_{n\times n}=A^T$是正定矩阵，对实$n$维线性空间$R^n$中的任意两个向量

$$\alpha=(x_1,x_2,\ \cdots,\ x_n),\ \beta=(y_1,\ y_2,\ \cdots,\ y_n)$$

规定　　$(\alpha,\ \beta)=\alpha A\beta^T$

(1) 证明：$(\alpha,\ \beta)$ 是内积，从而R^n对这个内积构成欧氏空间；

(2) 求单位向量$e_1=(1,\ 0,\ \cdots,\ 0)$，$e_2=(0,\ 1,\ 0,\ \cdots,\ 0)$，$\cdots$，$e_n=(0,\ 0,\ \cdots,\ 0,\ 1)$ 的度量矩阵；

(3) 写出在这个欧氏空间R^n中的Cauchy-Schwarz不等式。

5. 在欧氏空间R^4中求一单位向量与

$$\alpha=(1,\ \ 1,\ -1,\ 1)$$
$$\beta=(1,\ -1,\ -1,\ 1)$$
$$\gamma=(2,\ \ 1,\ \ 1,\ 3)$$

正交。

6. 设α_1，α_2，$\cdots$，α_n是欧氏空间V中的一组基底。证明：

(1) 若$\gamma\in V$使得 $(\gamma,\ \alpha_i)=0$，$i=1,\ 2,\ \cdots,\ n$

则$\gamma=0$。

(2) 若γ_1，$\gamma_2\in V$，使对任一$\beta\in V$，有

$$(\gamma_1,\ \beta)=(\gamma_2,\ \beta)$$

则　$\gamma_1=\gamma_2$

7. 设e_1，e_2，e_3是三维欧氏空间R^3中一组标准正交基底.证明：

$$\alpha_1=\frac{1}{3}(2e_1+2e_2-e_3)$$

$$\alpha_2=\frac{1}{3}(2e_1-e_2+2e_3)$$

$$\alpha_3=\frac{1}{3}(e_1-2e_2-2e_3)$$

也是一组标准正交基底。

8. 设α，β是欧氏空间R^n中任意两个向量，A是任意实n阶矩阵，证明：

$$(A\alpha, \beta)=(\alpha, A^T\beta)$$

9. 证明：n阶λ-矩阵

$$A(\lambda)=\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & a_n \\ -1 & \lambda & 0 & \cdots & 0 & a_{n-1} \\ 0 & -1 & \lambda & \cdots & 0 & a_{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & a_2 \\ 0 & 0 & 0 & \cdots & -1 & \lambda+a_1 \end{pmatrix}$$

非常数的不变因子只有一个

$$d_n(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n。$$

10. 设λ-矩阵

$$A(\lambda)=\begin{pmatrix} 3\lambda^2+2\lambda-3 & 2\lambda-1 & \lambda^2+2\lambda-3 \\ 4\lambda^2+3\lambda-5 & 3\lambda-2 & \lambda^2+3\lambda-4 \\ \lambda^2+\lambda-4 & \lambda-2 & \lambda-1 \end{pmatrix}$$

求$A(\lambda)$的不变因子和初等因子。

11. 求下列矩阵的Jordan标准形J。

$$A=\begin{pmatrix} 3 & 7 & -3 \\ -2 & -5 & 2 \\ -4 & -10 & 3 \end{pmatrix}, \quad B=\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -1 & 1 \\ 0 & 0 & 0 & 2 & 2 & -1 \\ 0 & 0 & 0 & 1 & 2 & -1 \end{pmatrix}$$

12. 设

$$A=\begin{pmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & 1 & 2 \end{pmatrix}$$

求A的Jordan标准形J，并求相似变换矩阵P，使 $P^{-1}AP=J$

第二章　矩阵多项式

第一节　多　项　式

在《高等代数》中已讲过数域F上的一元多项式：

$$f(x)=a_0+a_1x+\cdots+a_nx^n=\sum_{i=0}^{n}a_ix^i \qquad (2\text{-}1)$$

其中$a_i\in F(i=0,\ 1,\ \cdots,\ n)$，$x$为变量。

若$a_n\neq 0$，则n是多项式$f(x)$的次数(degree)。若$a_n=1$，则多项式就是上章所讲过的首一多项式。

我们知道，两个多项式的和、积，其定义如下：

对于两个多项式：$f(x)=\sum\limits_{i=0}^{n}a_ix^i$，$g(x)=\sum\limits_{i=0}^{n}b_ix^i$

其中a_i，$b_i\in F$ $(i=0,\ 1,\ \cdots,\ n)$。定义其和为：

$$f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i \qquad (2\text{-}2)$$

其积为

$$f(x)g(x)=\sum_{i=0}^{2n}c_ix^i \qquad (2\text{-}3)$$

其中　$c_i=\sum\limits_{\substack{j+k=i\\ j,\ k\leq n}}a_jb_k$，$c_i\in F$ $(i=0,\ 1,\ 2,\ \cdots,\ 2n)$

并且次数之间有下列关系：

$$\deg(f+g)\leq\max(\deg f,\ \deg g) \qquad (2\text{-}4)$$

$$\deg(fg)=\deg f+\deg g \qquad (2\text{-}5)$$

对多项式f，$g(g\neq0)$可选取适当的多项式q，r

使得 $$f=gq+r,\ \deg r<\deg g \tag{2-6}$$

其中q称为f除以g的商，r称为f的余式。当$r=0$,即f没有余式时，称g能除尽（整除）f，记为$g|f$。当f_1,f_2都可用g除尽，而且不能用比g次数高的任一多项式除尽时，称g是f_1,f_2的最大公因式。此时，若$g(x)$为首一多项式，则用$(f_1(x),\ f_2(x))$来表示。当$(f_1)x),f_2(x))=1$时，称f_1和f_2是互素的（或互质的）。

两个多项式f_1和f_2互素的充分必要条件是存在多项式$P_1(x)$，$P_2(x)$使得：

$$f_1P_1+f_2P_2=1 \tag{2-7}$$

使$f(x)=0$的x值，叫做$f(x)$的零点或根。

特别地，若数域F为复数域C，则$f(x)$具有n个根$x_1,x_2,\cdots,x_n\in C$。此时，$f(x)$可表示为

$$f(x)=a_n(x-x_1)(x-x_2)\cdots(x-x_n) \tag{2-8}$$

此式称为$f(x)$的质因子分解。

第二节　矩阵多项式

设A是一个$n\times n$阶方阵，$f(\lambda)$为数域F上的多项式。在$f(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_m\lambda^m$中把$\lambda^i$换成$A^i$；$\lambda^0=1$换成单位矩阵$I$，就得出一个矩阵多项式

$$f(A)=a_0I+a_1A+a_2A^2+\cdots+a_mA^m \tag{2-9}$$

称为矩阵A的多项式。若$a_m\neq0$,则m称为$f(A)$的次数。所以$f(A)$之“值”也是一个$n\times n$阶方阵。

例如：$f(A)=4I-3A+A^2+A^3$，其中 $A=\begin{pmatrix}1&1\\0&2\end{pmatrix}$

将 $A^2=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}=\begin{pmatrix}1 & 3\\ 0 & 4\end{pmatrix}$

$$A^3=\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}^2\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}=\begin{pmatrix}1 & 7\\ 0 & 8\end{pmatrix}$$

代入$f(A)$的表达式，得

$$f(A)=4\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-3\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}+\begin{pmatrix}1 & 3\\ 0 & 4\end{pmatrix}+\begin{pmatrix}1 & 7\\ 0 & 8\end{pmatrix}=\begin{pmatrix}3 & 7\\ 0 & 10\end{pmatrix}$$

两个矩阵多项式的和、积可根据（2-2）式和（2-3）式分别定义为：

$$f(A)+g(A)=(f+g)A \tag{2-10}$$

$$f(A)g(A)=(fg)A \tag{2-11}$$

记号$(f+g)A$与$(fg)A$分别表示把$f(\lambda)+g(\lambda)$与$f(\lambda)g(\lambda)$中的λ^i换成A^i，$\lambda^0=1$换成I。据此，由(2-11)式可得

$$f(A)g(A)=(fg)A=(gf)A=g(A)f(A)$$

即$f(A)$，$g(A)$可以交换，与矩阵乘法运算不可交换无关，于是，矩阵多项式有下列性质：

性质1 对任意矩阵A，矩阵多项式的乘法运算是可交换的。

性质2 $$f(PAP^{-1})=Pf(A)P^{-1} \tag{2-12}$$

其中 矩阵P是可逆的。

证明 $I=PIP^{-1}$，显然，$(PAP^{-1})^k=PA^kP^{-1}$（k为正整数），把这些式子代入$f(PAP^{-1})$中，再利用分配律，即可得到

$$f(PAP^{-1})=Pf(A)P^{-1}$$

注意：矩阵多项式与多项式矩阵不同，一定不要混淆。

因为多项式矩阵，是矩阵中以多项式f_{ij}为元素的$m\times n$阶矩阵$F(\lambda)=[f_{ij}(\lambda)]_{m\times n}$

设$p=\max\ \deg f_{ij}$，利用矩阵序列F_i $(i=0,\ 1,\ \cdots,p)$，则有

$$F(\lambda)=F_0+\lambda F_1+\lambda^2F_2+\cdots+\lambda^pF_p$$

显然与（2-9)式不同。

例如：

$$F(\lambda)=\begin{pmatrix}\lambda^3+2\lambda^2+\lambda-3 & \lambda-1\\ 0.2\lambda^2-\lambda+1.3 & \lambda^2-4.2\lambda-0.7\end{pmatrix}$$

$$=\lambda^3\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}+\lambda^2\begin{pmatrix}2 & 0\\ 0.2 & 1\end{pmatrix}$$

$$+\lambda\begin{pmatrix}1 & 1\\ -1 & -4.2\end{pmatrix}+\begin{pmatrix}-3 & -1\\ 1.3 & -0.1\end{pmatrix}$$

这种矩阵就是λ-矩阵。

定义 设A是$n\times n$阶方阵，若A满足

$$f(A)=0\quad(0\text{表示}n\times n\text{阶零方阵})$$

则多项式$f(\lambda)$称为A的**零化多项式**（Annihilating Polynomial)，又说$f(\lambda)$使A零化。

下面我们证明矩阵A的特征多项式为A的零化多项式，即凯莱-哈密尔顿（Cayley-Hamilton）定理。为此，先看一个引理

引理 设$f(\lambda)$为λ的多项式，A为n阶数量方阵，则存在n阶λ-矩阵$Q(\lambda)$，使

$$(\lambda I-A)Q(\lambda)=f(\lambda)I$$

或 $Q(\lambda)(\lambda I-A)=f(\lambda)I$

成立的充要条件是$f(A)=0$。（证明从略）

定理 （凯莱-哈密尔顿） 设$\psi(\lambda)=\det(\lambda I-A)$

则 $\psi(A)=0$

成立（其中A为n阶数量矩阵）。

证明 由于$\lambda I-A$的伴随矩阵 $(\lambda I-A)^*$也是λ-矩阵，且有

$$(\lambda I-A)(\lambda I-A)^*=\psi(\lambda)I \qquad (2\text{-}13)$$

从上述引理即知有$\psi(A)=0$

从定理可看到，n次多项式$\psi(\lambda)$是A的零化多项式。此外，次数比A的阶数低的多项式也可能使A零化。

例如：设 $A=\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}$

它的特征多项式为：$\psi(\lambda)=(\lambda-1)(\lambda-2)^2$，由上面定理知$\psi(A)=0$。但是，对于

$\psi_1(\lambda)=(\lambda-1)(\lambda-2)$也能得到$\psi_1(A)=0$，因为

$$\psi_1(A)=\left(\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}-\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$$
$$\left(\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}-\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right)$$
$$=\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

$$=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

又 $\psi_2(\lambda)=k(\lambda-1)(\lambda-2)$也可得到$\psi_2(A)=0$，这说明了使$A$零化的多项式不是唯一的。因此提出最小多项式的概念。

第三节 矩阵的最小多项式

定义1 在$n\times n$阶矩阵A的所有零化多项式中，次数最低的首一多项式，称为矩阵A的**最小多项式**（Minimal Polynomial），记为$\varphi_m(\lambda)$。由第二节中的定理，可知n阶矩阵A的最小多项式是存在的，且次数$\leqslant n$

例如：$A=\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

对于多项式$\psi(\lambda)=\lambda-2$有

$$\psi(A)=\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}-2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

所以$\psi(\lambda)=\lambda-2$是A的零化多项式。次数比$\psi(\lambda)$低的多项式是零次的。对首一零次多项式$\psi_1(\lambda)$，只有$\psi_1(\lambda)=1$，而

$$\psi_1(A)=I=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\neq\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

因此$\psi_1(\lambda)$不是A的零化多项式。

所以$\psi(\lambda)=\lambda-2=\varphi_m(\lambda)$是$A$的最小多项式。

定理1 n阶矩阵A的最小多项式是唯一的。

证明 如果$\varphi_m^1(\lambda)$，$\varphi_m^2(\lambda)$都是A的最小多项式，则其次数必须相等。若是$\varphi_m^1(\lambda)\neq\varphi_m^2(\lambda)$，命$\psi(\lambda)=\varphi_m^1(\lambda)-\varphi_m^2(\lambda)$，

其次数一定比$\varphi_m^1(\lambda)$，$\varphi_m^2(\lambda)$的次数低，设$\psi(\lambda)$的首项系数为a，则$\psi_1(\lambda)=\frac{1}{a}\psi(\lambda)$ 是首一多项式，且$\psi_1(A)=\frac{1}{a}\psi(A)$

$=\frac{1}{a}(\varphi_m^1(A)-\varphi_m^2(A))=0$ 即$\psi_1(\lambda)$是A的零化多项式，而且次数又比$\varphi_m^1(\lambda)$、$\varphi_m^2(\lambda)$的次数低，这与$\varphi_m^1(\lambda)$、$\varphi_m^2(\lambda)$是A的最小多项式的假设矛盾。因此，必有$\varphi_m^1(\lambda)\equiv\varphi_m^2(\lambda)$。故$A$的最小多项式是唯一的。

定理2 n阶矩阵A的最小多项式能整除A的任意零化多项式。

证明 设$\varphi_m(\lambda)$是A的最小多项式，对于$\deg f(\lambda)>\deg\varphi_m(\lambda)$的$A$的任意零化多项式$f(\lambda)$，可表示为

$f(\lambda)=\varphi_m(\lambda)g(\lambda)+r(\lambda)$，$\deg r(\lambda)<\deg\varphi_m(\lambda)$，如果$\varphi_m(\lambda)$不能整除$f(\lambda)$，则$r(\lambda)\neq 0$，但$f(\lambda)$是$A$的零化多项式，即有$f(A)=0$。又因$\varphi_m(A)=0$，于是$r(A)=0$必成立，而$r(\lambda)$的次数又小于$\varphi_m(\lambda)$的次数，这与$\varphi_m(\lambda)$是$A$的最小多项式相矛盾，所以必有$r(\lambda)=0$，故$\varphi_m(\lambda)$能整除$f(\lambda)$。证毕。

推论 设A的特征多项式为$\psi(\lambda)$，则$\varphi_m(\lambda)\mid\psi(\lambda)$。

定理3 n阶矩阵A的最小多项式为

$$\varphi_m(\lambda)=\frac{\psi(\lambda)}{D_{n-1}(\lambda)} \tag{2-14}$$

其中$\psi(\lambda)=|\lambda I-A|$，$D_{n-1}(\lambda)$是$\lambda I-A$所有$n-1$阶子式的最大公因式（即$n-1$阶行列式因子）。

证明 记 $\varphi(\lambda)=\dfrac{\psi(\lambda)}{D_{n-1}(\lambda)}$

现要证明$\varphi(\lambda)$就是A的最小多项式$\varphi_m(\lambda)$。因为$D_{n-1}(\lambda)$能除尽 $(\lambda I-A)^*$的每个元素，故可写

$$(\lambda I-A)^*=D_{n-1}(\lambda)C(\lambda)$$

其中$C(\lambda)$为各元素互质的λ-矩阵，由于

$$(\lambda I-A)(\lambda I-A)^*=\psi(\lambda)I$$

故有

$$(\lambda I-A)C(\lambda)=\varphi(\lambda)I \tag{2-15}$$

根据前节引理，即知有$\varphi(A)=0$，即$\varphi(\lambda)$是矩阵A的零化多项式。

今再证$\varphi(\lambda)$是矩阵A的最小多项式，由定理2，有$\varphi_m(\lambda)/\varphi(\lambda)$，于是存在多项式$X(\lambda)$，使

$$\varphi(\lambda)=X(\lambda)\varphi_m(\lambda) \tag{2-16}$$

因为$\varphi_m(A)=0$，根据前节引理，又存在一个λ-矩阵$Q(\lambda)$，使

$$(\lambda I-A)Q(\lambda)=\varphi_m(\lambda)I \tag{2-17}$$

由（2-16）式，有

$$\varphi(\lambda)I=X(\lambda)\varphi_m(\lambda)I=(\lambda I-A)X(\lambda)Q(\lambda)$$

再由（2-15）式，便可推出

$$C(\lambda)=X(\lambda)Q(\lambda)$$

此式表明$X(\lambda)$是$C(\lambda)$中各元素的公因式，故有$X(\lambda)\equiv 1$。这样，由（2-16）式，即知$\varphi(\lambda)=\varphi_m(\lambda)$。证毕。

推论 矩阵A的特征矩阵$\lambda I-A$的第n个不变因子$d_n(\lambda)$就是A的最小多项式$\varphi_m(\lambda)$，即

$$\varphi_m(\lambda)=d_n(\lambda) \tag{2-18}$$

证明 因为$\psi(\lambda)=|\lambda I-A|=D_n(\lambda)$，故有

$$\varphi_m(\lambda)=\frac{\psi(\lambda)}{D_{n-1}(\lambda)}=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}=d_n(\lambda) \tag{2-19}$$

定理4 n阶矩阵A的特征多项式的根，是A的最小多项

式的根。

证明 将A的特征矩阵$\lambda I-A$化为Smith标准形。即：

$$\lambda I-A\cong\begin{pmatrix}d_1(\lambda) & & & & \\ & d_2(\lambda) & & & \\ & & \ddots & & \\ & & & d_{n-1}(\lambda) & \\ & & & & d_n(\lambda)\end{pmatrix}$$

则有$\psi(\lambda)=|\lambda I-A|=cd_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)$①（其中$c$是不等于零的常数。实际上，这时$c=1$）。若$\lambda_1$是$\psi(\lambda)=0$的一个根，即有$\psi(\lambda_1)=0$。由此知$\lambda_1$必为某$d_j(\lambda)$之根。又$d_j(\lambda)|d_n(\lambda)$，即得$d_n(\lambda_1)=0$，由上推论知，有$\varphi_m(\lambda_1)=0$故有$\varphi_m(\lambda)=0$证毕。

根据此定理，可知，若

$$\psi(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_t)^{k_t}$$

其中$k_1+k_2+\cdots+k_t=n$（A的阶数），$k_i>0$（$i=1, 2, \cdots, t$）

且当$i\neq j$时，$\lambda_i\neq\lambda_j$

则 $\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t}$

其中 $1\leqslant m_i\leqslant k_i$（$i=1, 2, \cdots, t$）

特别地，当$k_i=1$时，$t=n$，必有$m_i=1$ $i=1,2,\cdots,t$

则$\varphi_m(\lambda)=\psi(\lambda)$ 即得如下结论：

推论 矩阵A的n个特征根互异，则A的最小多项式就等于A的特征多项式。

定理5 设n阶矩阵A的最小多项式为

$$\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t} \quad (2-20)$$

---

①两个等价的λ-矩阵的行列式，只相差一个非零的常数因子

其中λ_1，λ_2，…，λ_t是A的相异特征值，则对于两个多项式$f(\lambda)$，$g(\lambda)$使得$f(A)=g(A)$的充分必要条件为：

$$f^{(l)}(\lambda_j)=g^{(l)}(\lambda_j),\quad j=1,2,\cdots,t;\quad l=0,1,2,\cdots,m_j-1 \tag{2-21}$$
②

证明 必要性：

若 $f(A)=g(A)$，则 $f(A)-g(A)=0$ 于是 $h(\lambda)=f(\lambda)-g(\lambda)$ 是A的零化多项式，根据定理2，则$h(\lambda)$能用最小多项式$\varphi_m(\lambda)$整除，即存在多项式$q(\lambda)$，使得

$h(\lambda)=q(\lambda)\varphi_m(\lambda)$ 由 (2-20)式，显然

$h^{(l)}(\lambda_j)=0$，$j=1,2,\cdots,t$；$l=0,1,2,\cdots,m_j-1$，

故 (2-21) 式成立。

充分性：

若 (2-21) 式成立，则$h(\lambda)=f(\lambda)-g(\lambda)$ 能用$\varphi_m(\lambda)$整除，因此$h(A)=0$. 故$f(A)=g(A)$成立。证毕

定义2 设n阶矩阵A的最小多项式为

$$\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t}$$

若函数$f(\lambda)$在A的相异特征值λ_1，λ_2，…，λ_t处都有下列k个确定的有限值 ($k=m_1+m_2+\cdots+m_t$)

$$f^{(l)}(\lambda_j)=\left.\frac{\mathrm{d}^{(l)}f(\lambda)}{\mathrm{d}\lambda^{(l)}}\right|_{\lambda=\lambda_j},\qquad j=1,2,\cdots,t;$$

$$l=0,1,2,\cdots,m_j-1$$

则称函数$f(\lambda)$定义在A的谱上。这些值称为$f(\lambda)$在A上的谱值。

---

② $f^{(l)}(\lambda_j)=\left[\frac{\mathrm{d}^{(l)}f(\lambda)}{\mathrm{d}\lambda^{(l)}}\right]_{\lambda=\lambda_j}$

从而（2-21）式左，右端的值又叫做多项式$f(\lambda)$和$g(\lambda)$在矩阵A上的谱值。于是有

推论1　对于某多项式$f(\lambda)$，$f(A)=0$的充分必要条件是$f(\lambda)$在矩阵A上的谱值为零。即是

$$f^{(l)}(\lambda_j)=0,\quad j=1,\ 2,\ \cdots,\ t;$$
$$l=0,\ 1,\ 2,\ \cdots,\ m_j-1$$

推论2　对于两个多项式$f(\lambda)$和$g(\lambda)$，A是已知矩阵，则$f(A)=g(A)$的充分必要条件是：多项式$f(\lambda)$和$g(\lambda)$在矩阵A上的谱值相同。

习　题　二

1．试证任意矩阵A的逆矩阵A^{-1}，可以表为矩阵A的多项式。

2．设 $A=\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix}$ 和 $B=\begin{pmatrix} 1 & 2 & 2 \\ 1 & -1 & 1 \\ 4 & -12 & 1 \end{pmatrix}$

试计算：

(1)　$(A^4-4A^3+6A^2+6A-8I)^{-1}$

(2)　$f(B)=B^7-B^5-19B^4+28B^3+6B-4I$

3．求矩阵

$$A=\begin{pmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{pmatrix} \text{ 及 } B=\begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ -a_1 & a_0 & -a_3 & a_2 \\ -a_2 & a_3 & a_0 & -a_1 \\ -a_3 & -a_2 & a_1 & a_0 \end{pmatrix}$$

的最小多项式。

4．证明：

(1) 相似矩阵有相同的最小多项式，逆定理不成立。

(2) 任意矩阵与它的转置矩阵有相同的最小多项式。

第三章 矩 阵 分 析

在一般的线性代数中，主要是用代数的方法来研究矩阵，没有涉及到极限和微积分的运算规律。由于科学的不断发展，矩阵理论在各个学科领域中的应用愈来愈广泛。因此，我们有必要对矩阵分析作些介绍。

矩阵分析理论也与数学分析一样，是建立在极限概念之上。因此，本章首先讨论向量和矩阵的极限；又因在高维空间讨论数值逼近和研究微分方程的数值解时，常常需要考虑逼近程度的问题，这就涉及向量和矩阵的范数概念，所以本章还将讨论向量和矩阵的范数。

第一节 向量和矩阵的极限

一、向量的极限

定义1 对于向量序列$\{X^{(k)}\}$，

其中 $X^{(k)}=(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$

若其每一分量$x_i^{(k)}$当$k\to\infty$时，都有极限x_i，

即 $\lim\limits_{k\to\infty} x_i^{(k)}=x_i, \quad i=1, 2, \cdots, n$ (3-1)

则称向量序列$\{X^{(k)}\}$有极限$X=(x_1, x_2, \cdots, x_n)$，或称$\{X^{(k)}\}$收敛于$X$，记为$\lim\limits_{k\to\infty} X^{(k)}=X$， (3-2)

或 $X^{(k)}\to X(k\to\infty)$

例如：在R^2中的向量序列

$$\begin{pmatrix}2.1\\0.9\end{pmatrix}, \begin{pmatrix}2.01\\0.99\end{pmatrix}, \begin{pmatrix}2.001\\0.999\end{pmatrix}, \cdots,$$

$\begin{pmatrix}2+10^{-k}\\1-10^{-k}\end{pmatrix}$…收敛于向量$\begin{pmatrix}2\\1\end{pmatrix}$

又如 $X^{(k)}=\begin{pmatrix}\dfrac{1}{2^k}\\ \dfrac{\sin k}{k}\end{pmatrix}$，有 $\lim\limits_{k\to\infty}\dfrac{1}{2^k}=0$，$\lim\limits_{k\to\infty}\dfrac{\sin k}{k}=0$，

故 $X^{(k)}\to\begin{pmatrix}0\\0\end{pmatrix}\ (k\to\infty)$

但向量序列 $\begin{pmatrix}1\\2\end{pmatrix}$，$\begin{pmatrix}1\\-2\end{pmatrix}$，$\begin{pmatrix}1\\2\end{pmatrix}$，$\begin{pmatrix}1\\-2\end{pmatrix}$，

…，$\begin{pmatrix}1\\(-1)^{k+1}\ 2\end{pmatrix}$，…没有极限。因为第二个分量没有极限。

由于实数序列的极限是唯一的，所以定义 1 中向量序列的极限必定是唯一的。

显然，$\lim\limits_{k\to\infty}X^{(k)}=X \iff \lim\limits_{k\to\infty}(X^{(k)}-X)=0$

即向量序列$\{X^{(k)}\}$收敛于X的充要条件是向量序列$\{X^{(k)}-X\}$收敛于零向量。

根据向量的极限，可引出向量级数的极限。

设向量级数

$$X^{(1)}+X^{(2)}+\cdots+X^{(k)}+\cdots$$

其部分和为：$Y^{(k)}=X^{(1)}+X^{(2)}+\cdots+X^{(k)}$。如果向量序列$\{Y^{(k)}\}$有极限，则称此级数有极限（或收敛）。$Y^{(k)}$的极限就是此级数的极限（或称为此级数的和）。反之，若向量级数没有极限，就称此级数是发散的。

例如：$X^{(k)}=\begin{pmatrix}\dfrac{1}{2^k}\\ \dfrac{1}{k}\end{pmatrix}$；

$$Y^{(k)}=X^{(1)}+X^{(2)}+\cdots+X^{(k)}=\begin{pmatrix}\sum_{i=1}^{k}\dfrac{1}{2^i}\\ \sum_{i=1}^{k}\dfrac{1}{i}\end{pmatrix}.$$

因为 $\sum_{i=1}^{k}\dfrac{1}{2^i}=\dfrac{1}{2}\dfrac{1-(\frac{1}{2})^k}{1-\frac{1}{2}}\to 1\quad(k\to\infty)$

而 $\sum_{i=1}^{k}\dfrac{1}{i}\to\infty\quad(k\to\infty)$

则$Y^{(k)}$的极限不存在，故向量级数

$$\sum_{k=1}^{\infty}X^{(k)}=X^{(1)}+X^{(2)}+\cdots+X^{(k)}+\cdots$$

发散。

二、矩阵序列的极限

矩阵序列的收敛性和向量类似。

定义2 设有同阶矩阵序列$\{A_k\}=\{[a_{ij}^{(k)}]_{m\times n}\}$，若存在矩阵$A=[a_{ij}]_{m\times n}$当$k\to\infty$时，$a_{ij}^{(k)}$收敛于$a_{ij}$。即

$$\lim_{k\to\infty}a_{ij}^{(k)}=a_{ij}\quad(i=1,\ 2,\ \cdots,\ m;\ j=1,\ 2,\ \cdots,\ n)\tag{3-3}$$

称矩阵序列$\{A_k\}$收敛于A。即

$$\lim_{k\to\infty} A_k = A \tag{3-4}$$

矩阵A称为矩阵序列$\{A_k\}$的极限。

不收敛的矩阵序列统称为发散的。

收敛的矩阵序列与收敛的数列有类似的性质：

性质1 对于矩阵序列$\{A_k\}$和$\{B_k\}$，若当$k\to\infty$时，有$A_k\to A$，$B_k\to B$，则必有$\alpha A_k+\beta B_k\to\alpha A+\beta B$

即
$$\lim_{k\to\infty}(\alpha A_k+\beta B_k)=\alpha A+\beta B \tag{3-5}$$

其中α，β为两个任意常数。

性质2 若当$k\to\infty$时，有$A_k\to A$和$B_k\to B$，则必有$A_kB_k\to AB$ 即

$$\lim_{k\to\infty}A_kB_k=AB \tag{3-6}$$

性质3 若$\{A_k\}$收敛于A，当逆矩阵A_k^{-1}，A^{-1}均存在时，则必有$\{A_k^{-1}\}$收敛于A^{-1}。即

$$\lim_{n\to\infty}A_k^{-1}=A^{-1} \tag{3-7}$$

证明 对性质1由定义2可直接得知。只需证明性质2和性质3。

证明性质2： 因为$(A_kB_k)_{ij}=\sum_{t=1}^{n}(A_k)_{it}(B_k)_{tj}$

$$(AB)_{ij}=\sum_{t=1}^{n}A_{it}B_{tj}$$

由于当$k\to\infty$时 有$A_k\to A$，$B_k\to B$，故有

$$\sum_{t=1}^{n}(A_k)_{it}(B_k)_{tj}\to\sum_{t=1}^{n}A_{it}B_{tj} \qquad \text{即}$$

$$(A_kB_k)_{ij}\to(AB)_{ij}$$

所以
$$\lim_{k\to\infty}A_kB_k=AB$$

须指出，若是矩阵序列不是方阵时，只要$\{A_k\}$是$s\times p$阶

矩阵序列和$\{B_k\}$是$p\times r$阶矩阵序列，性质2仍然成立。

证明性质3：因为$A_k^{-1}=A_k^*/|A_k|$，而且$|A_k|$和A_k^*中各元素都是A_k中的元素的多项式。

由于$\lim\limits_{k\to\infty}$多项式（A_k的元素）=多项式($\lim\limits_{k\to\infty}A_k$的元素)

则 $\lim\limits_{k\to\infty}A_k^*=A^*$，$\lim\limits_{k\to\infty}|A_k|=|A|\neq 0$

故 $\lim\limits_{k\to\infty}A_k^{-1}=A^*/|A|=A^{-1}$　　证毕

三、矩阵级数

控制系统理论中常遇到的e^A是作为无穷级数的和定义的。所以矩阵级数在系统理论中有重要的应用。

定义3　由矩阵序列$\{A_k\}$组成的无穷级数

即
$$\sum_{k=1}^{\infty}A_k=A_1+A_2+\cdots+A_k+\cdots \tag{3-8}$$

称为矩阵级数。其部分和记为

$$S_k=A_1+A_2+\cdots+A_k,$$

当$k\to\infty$时，如果$\{S_k\}$的极限存在，且等于S，

即 $\lim\limits_{n\to\infty}S_k=S$

则称矩阵级数$\sum\limits_{k=1}^{\infty}A_k$是收敛的，且矩阵$S$称为矩阵级数$\sum\limits_{n=1}^{\infty}A_k$的“**和矩阵**”。

若设$A_k=[a_{ij}^{(k)}]_{n\times n}$（$k=1, 2, \cdots,$ ），则可以看出，当且仅当由其元素组成的n^2个数项级数

$$\sum_{k=1}^{\infty}a_{ij}^{(k)}=a_{ij}^{(1)}+a_{ij}^{(2)}+\cdots+a_{ij}^{(k)}+\cdots$$

$$(i,\ j=1,\ 2,\ \cdots,\ n)$$

收敛时，矩阵级数$\sum_{k=1}^{\infty} A_k$才收敛，否则发散。

于是，矩阵级数的运算法则与数项级数的运算法则就完全一致。

四、函数矩阵

定义4 设矩阵

$$A(x)=\begin{pmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{pmatrix} \tag{3-9}$$

其中每个元素都是实变量x的函数（或复变量函数），则称矩阵$A(x)$为**函数矩阵。**

若对每一个元素$a_{ij}(x)$，当$x\to x_0$时，都存在着极限a_{ij}（常数），即

$$\lim_{x\to x_0} a_{ij}(x)=a_{ij}\quad (i=1,\ 2,\ \cdots,\ m;\ j=1,\ 2,\ \cdots,\ n)$$

则称矩阵$A(x)$有极限，且极限值为矩阵$A=[a_{ij}]_{m\times n}$

即

$$\lim_{x\to x_0} A(x)=\begin{pmatrix} \lim\limits_{x\to x_0} a_{11}(x) & \lim\limits_{x\to x_0} a_{12}(x) & \cdots & \lim\limits_{x\to x_0} a_{1n}(x) \\ \lim\limits_{x\to x_0} a_{21}(x) & \lim\limits_{x\to x_0} a_{22}(x) & \cdots & \lim\limits_{x\to x_0} a_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ \lim\limits_{x\to x_0} a_{m1}(x) & \lim\limits_{x\to x_0} a_{m2}(x) & \cdots & \lim\limits_{x\to x_0} a_{mn}(x) \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}=A \tag{3-10}$$

函数矩阵的极限，具有函数极限的类似性质。

例如：当$x \to x_0$时，设函数矩阵$A(x)$和$B(x)$有极限A和B，则有

$$
\begin{aligned}
&(1)\quad \lim_{x \to x_0}[A(x)+B(x)]=A+B \\
&(2)\quad \lim_{x \to x_0}[A(x)B(x)]=AB \\
&(3)\quad \lim_{x \to x_0} kA(x)=kA
\end{aligned}
\tag{3-11}
$$

其中矩阵A，B均为常数矩阵，k为常数。

同样可以给出函数矩阵在某一点处或某一区间上连续的定义。

定义5 若函数矩阵$A(x)=[a_{ij}(x)]_{m\times n}$中所有元素$a_{ij}(x)$（$i=1, 2, \cdots, m$；$j=1, 2, \cdots, n$）在某一点$x_0$处或某一区间上连续，则称函数矩阵$A(x)$在此点$x_0$处或此区间上连续。

类似地，可以推广到多变量的函数矩阵。例如：

$C(x, y)=[c_{ij}(x, y)]_{m\times n}$及$C(x_1, x_2, x_3, x_4)=[c_{ij}(x_1, x_2, x_3, x_4)]_{m\times n}$

也具有上述类似的定义和性质。

第二节 函数矩阵的微分和积分

本节所研究的内容：是用矩阵来描述微积分中的若干结果，它在实用上很方便。我们将着重讨论在工程实际中常见的三个问题，即是函数矩阵关于自变量的微分和积分；纯量函数关于矩阵的微分；向量函数关于向量的微分。

一、函数矩阵的微分和积分

定义1 设函数矩阵$A(x)=[a_{ij}(x)]_{m\times n}$（$i=1, 2, \cdots, m$；$j=1, 2, \cdots, n$），如果对于所有元素$a_{ij}(x)$在$x=x_0$点或某一区间上是可微的，则称该函数矩阵$A(x)$在$x=x_0$点或

某一区间上是可微的。并称

$$\frac{\mathrm{d}}{\mathrm{d}x}A(A)=A'(x)=\begin{pmatrix} a'_{11}(x) & a'_{12}(x) & \cdots & a'_{1n}(x) \\ a'_{21}(x) & a'_{22}(x) & \cdots & a'_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ a'_{m1}(x) & a'_{m2}(x) & & a'_{mn}(x) \end{pmatrix} \tag{3-12}$$

为函数矩阵$A(x)$对x的导数。

显然，此导数也可以用数学分析中数量函数的导数定义的形式来定义。即

$$\frac{\mathrm{d}}{\mathrm{d}x}A(x)=\lim_{\Delta x\to 0}\frac{A(x+\Delta x)-A(x)}{\Delta x}$$

$$=\lim_{\Delta x\to 0}\left(\frac{a_{ij}(x+\Delta x)-a_{ij}(x)}{\Delta x}\right)_{m\times n}$$

由函数矩阵的导数定义，可以得到关于函数矩阵的导数运算法则。

设函数矩阵$A(x)$与$B(x)$是可微的，则有

法则1　$A(x)+B(x)$也是可微的，且满足

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)+B(x)]=\frac{\mathrm{d}}{\mathrm{d}x}A(x)+\frac{\mathrm{d}}{\mathrm{d}x}B(x) \tag{3-13}$$

法则2　$$\frac{\mathrm{d}}{\mathrm{d}x}[kA(x)]=k\frac{\mathrm{d}}{\mathrm{d}x}A(x) \tag{3-14}$$

其中k是任意常数。

一般地，若K是一个常数矩阵，则有

$$\frac{\mathrm{d}}{\mathrm{d}x}[KA(x)]=K\frac{\mathrm{d}}{\mathrm{d}x}A(x) \tag{3-15}$$

或

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)K]=\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)\cdot K \tag{3-16}$$

以上两个法则请读者自己证明。

法则3 $A(x)B(x)$也是可微的，且有

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)B(x)]=\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)B(x)+A(x)\frac{\mathrm{d}}{\mathrm{d}x}B(x) \tag{3-17}$$

证明 由导数定义，有

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)B(x)]=\lim_{\Delta x\to 0}\frac{A(x+\Delta x)B(x+\Delta x)-A(x)B(x)}{\Delta x}$$

$$=\lim_{\Delta x\to 0}\left(\frac{A(x+\Delta x)B(x+\Delta x)-A(x)B(x+\Delta x)}{\Delta x}\right.$$

$$\left.+\frac{A(x)B(x+\Delta x)-A(x)B(x)}{\Delta x}\right)$$

$$=\lim_{\Delta x\to 0}\frac{A(x+\Delta x)-A(x)}{\Delta x}B(x+\Delta x)+\lim_{\Delta x\to 0}A(x)$$

$$\times\frac{B(x+\Delta x)-B(x)}{\Delta x}=\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)B(x)+A(x)\frac{\mathrm{d}}{\mathrm{d}x}B(x)$$

证毕。

这个公式可以推广到有限个函数矩阵的乘积。即

$$\frac{\mathrm{d}}{\mathrm{d}x}[A_1(x)A_2(x)\cdots A_n(x)]=\left(\frac{\mathrm{d}}{\mathrm{d}x}A_1(x)\right)A_2(x)\cdots A_n(x)$$

$$+A_1(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A_2(x)\right)A_3(x)\cdots A_n(x)+\cdots$$

$$+A_1(x)A_2(x)\cdots A_{n-1}(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A_n(x)\right)$$

其中$A_1(x)$，$A_2(x)$，…，$A_n(x)$都是可微的。

注意：上面导数公式中的乘积次序是不能交换的。例如，计算一个函数矩阵的平方的导数：

$$\frac{\mathrm{d}}{\mathrm{d}x}A^2(x)=\frac{\mathrm{d}}{\mathrm{d}x}[A(x)A(x)]$$

$$=\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)A(x)+A(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)$$

不能写成$\frac{\mathrm{d}}{\mathrm{d}x}A^2(x)=2A(x)\frac{\mathrm{d}}{\mathrm{d}x}A(x)$

只有在$f(x)$表示一个纯量函数（即非函数矩阵）的情形时，式中各项的次序才可以交换。例如

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x)A(x)]=f'(x)A(x)+f(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)$$

或

$$\frac{\mathrm{d}}{\mathrm{d}x}[f(x)A(x)]=A(x)f'(x)+\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)f(x)$$

〔**例1**〕 求二次型X^TAX对变量t的导数。

其中

$$X=\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad A=[a_{ij}]_{n\times n}, \quad 且a_{ij}=a_{ji}$$

解 $$\frac{\mathrm{d}}{\mathrm{d}t}[X^TAX]=\frac{\mathrm{d}X^T}{\mathrm{d}t}(AX)+X^T\frac{\mathrm{d}}{\mathrm{d}t}(AX)$$

$$=\frac{\mathrm{d}X^T}{\mathrm{d}t}(AX)+X^T\frac{\mathrm{d}A}{\mathrm{d}t}X+X^TA\frac{\mathrm{d}X}{\mathrm{d}t}$$

根据二次型定义，可得

$$\frac{\mathrm{d}X^T}{\mathrm{d}t}(AX)=X^TA^T\frac{\mathrm{d}X}{\mathrm{d}t}=X^TA\frac{\mathrm{d}X}{\mathrm{d}t}$$

于是

$$\frac{\mathrm{d}}{\mathrm{d}t}[X^TAX]=2X^TA\frac{\mathrm{d}X}{\mathrm{d}t}+X^T\frac{\mathrm{d}A}{\mathrm{d}t}X$$

法则4 设矩阵A是u的函数，$u=f(x)$，则有

$$\frac{\mathrm{d}}{\mathrm{d}x}\{A[f(x)]\}=\frac{\mathrm{d}}{\mathrm{d}u}A(u)\cdot f'(x) \tag{3-18}$$

或

$$\frac{\mathrm{d}}{\mathrm{d}x}\{A[f(x)]\}=f'(x)\frac{\mathrm{d}}{\mathrm{d}u}A(u)$$

证明 $\frac{\mathrm{d}}{\mathrm{d}x}\{A[f(x)]\}=\frac{\mathrm{d}}{\mathrm{d}u}\{A[f(x)]\}\frac{\mathrm{d}u}{\mathrm{d}x}=\frac{\mathrm{d}}{\mathrm{d}u}A(u)f'(x)$

此法则与数学分析中复合函数求导法则基本一致。

法则5 若矩阵$A(x)$和其逆矩阵$A^{-1}(x)$都是可微的，则

$$\frac{\mathrm{d}}{\mathrm{d}x}A^{-1}(x)=-A^{-1}(x)\frac{\mathrm{d}A(x)}{\mathrm{d}x}A^{-1}(x) \tag{3-19}$$

证明 由逆矩阵定义，有

$$A(x)A^{-1}(x)=I$$

对上式两边求导：

$$\frac{\mathrm{d}}{\mathrm{d}x}[A(x)A^{-1}(x)]=\frac{\mathrm{d}}{\mathrm{d}x}I=0\text{（表示零矩阵）。}$$

利用（3-17）式，有

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(A(x)A^{-1}(x)\right)=\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)A^{-1}(x)+A(x)\frac{\mathrm{d}}{\mathrm{d}x}A^{-1}(x)$$

于是

$$A(x)\frac{\mathrm{d}}{\mathrm{d}x}A^{-1}(x)=-\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)A^{-1}(x)$$

故 $\frac{\mathrm{d}}{\mathrm{d}x}A^{-1}(x)=-A^{-1}(x)\left(\frac{\mathrm{d}}{\mathrm{d}x}A(x)\right)A^{-1}(x)$ 证毕。

由于（3-12）式定义的$\frac{d}{dx}A(x)$仍然是一个函数矩阵，所以当$a_{ij}(x)$（$i=1, 2, \cdots, m$；$j=1, 2, \cdots, n$）二阶可微时，也可以定义函数矩阵的二阶导数

$$\frac{d^2}{dx^2}A(x)=A''(x)=\lim_{\Delta x\to 0}\frac{A'(x+\Delta x)-A'(x)}{\Delta x}$$

而且，对更高阶导数也同样的有

$$\frac{d^{(n)}}{dx^{(n)}}A(x)=A^{(n)}(x)=\lim_{\Delta x\to 0}\frac{A^{(n-1)}(x+\Delta x)-A^{(n-1)}(x)}{\Delta x}$$

此外，对于多个自变量的函数矩阵的偏导数也是类似的定义。例如

设$A(x_1, x_2, x_3, x_4)=[a_{ij}(x_1, x_2, x_3, x_4)]_{m\times n}$

$$\frac{\partial}{\partial x_k}A(x_1,x_2,x_3,x_4)=\left(\frac{\partial}{\partial x_k}a_{ij}(x_1,x_2,x_3,x_4)\right)_{m\times n}$$

$k=1, 2, 3, 4$

下面再来讨论函数矩阵的积分问题。

定义2 设函数矩阵$A(x)=[a_{ij}(x)]_{m\times n}$，若$a_{ij}(x)$在区间$[a, b]$上对$x$可积，即积分

$\int_a^b a_{ij}(x)dx$存在，我们定义

$$\int_a^b A(x)dx=\begin{pmatrix}\int_a^b a_{11}(x)dx & \int_a^b a_{12}(x)dx & \cdots & \int_a^b a_{1n}(x)dx\\ \int_a^b a_{21}(x)dx & \int_a^b a_{22}(x)dx & \cdots & \int_a^b a_{2n}(x)dx\\ \cdots & \cdots & \cdots & \cdots\\ \int_a^b a_{m1}(x)dx & \int_a^b a_{m2}(x)dx & \cdots & \int_a^b a_{mn}(x)dx\end{pmatrix}_{m\times n}$$

称为$A(x)$在区间$[a, b]$上对x的积分，简记为

$$\int_a^b A(x)dx=\left(\int_a^b a_{ij}(x)dx\right)_{m\times n} \tag{3-20}$$

显然，对不定积分也可同样定义为

$$\int A(x)\mathrm{d}x=\left(\int a_{ij}(x)\mathrm{d}x\right)_{m\times n}$$

二、纯量函数关于矩阵的微分

设函数$f(X)$是以矩阵$X=[x_{ij}]_{m\times n}$中的$m\times n$个元素x_{ij}为自变量的可微纯量函数。即

$$f(X)=f(x_{11},\ x_{12},\ \cdots,\ x_{1n};\ x_{21},\ x_{22},\ \cdots,x_{2n};\ \cdots;\ x_{m1},\ x_{m2},\ \cdots,\ x_{mn})$$

则$f(X)$对矩阵X的导数有如下定义：

定义3 设矩阵$X=[x_{ij}]_{m\times n}$，若纯量函数$f(X)$对自变量x_{ij}可微，则定义f对矩阵X的导数如下

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1n}} \\ \frac{\partial f}{\partial x_{21}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{2n}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f}{\partial x_{m1}} & \frac{\partial f}{\partial x_{m2}} & \cdots & \frac{\partial f}{\partial x_{mn}} \end{pmatrix} \tag{3-21}$$

或

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\left(\frac{\partial f}{\partial x_{ij}}\right)_{m\times n}$$

〔**例2**〕 设矩阵

$$X=\begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix}$$

和纯量函数

$$f(X)=x_{11}^2+x_{12}^2+\cdots+x_{1n}^2+x_{21}^2+x_{22}^2+\cdots+x_{2n}^2+\cdots$$
$$+x_{m1}^2+x_{m2}^2+\cdots+x_{mn}^2$$

求$\frac{\mathrm{d}f}{\mathrm{d}X}$

解 因为

$$\frac{\partial f}{\partial x_{ij}}=2x_{ij}\ (i=1,\ 2,\ \cdots,\ m;\ j=1,\ 2,\ \cdots,\ n)$$

所以 $$\frac{\mathrm{d}}{\mathrm{d}X}f(X)=2\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}=2X$$

又因为有 $f(X)=\mathrm{tr}(XX^{\mathrm{T}})$

由此可得出关于迹（trace）的一些重要公式：

(1) $$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(XX^{\mathrm{T}})=2X \tag{3-22}$$

(2) $$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(BX)=\frac{\mathrm{d\,tr}(X^{\mathrm{T}}B^{\mathrm{T}})}{\mathrm{d}X}=B^{\mathrm{T}} \tag{3-23}$$

(3) $$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(X^{\mathrm{T}}AX)=(A+A^{\mathrm{T}})X \tag{3-24}$$

其中 $X=[x_{ij}]_{m\times n}$；$B=[b_{ij}]_{n\times m}$；$A=[a_{ij}]_{m\times m}$

证明 (2) 因为$B=[b_{ij}]_{n\times m}$是常量矩阵，$X=[x_{ij}]_{m\times n}$是变量矩阵。由矩阵乘法，可知

$$\mathrm{tr}(BX)=\mathrm{tr}(X^{\mathrm{T}}B^{\mathrm{T}})$$

$$=\sum_{i=1}^{m}b_{1i}x_{i1}+\sum_{i=1}^{m}b_{2i}x_{i2}+\cdots+\sum_{i=1}^{m}b_{ni}x_{in}$$

故 $\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(BX)=\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(X^{\mathrm{T}}B^{\mathrm{T}})$

$$=\begin{pmatrix} b_{11} & b_{21} & \cdots & b_{n1} \\ b_{12} & b_{22} & \cdots & b_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ b_{1m} & b_{2m} & \cdots & b_{nm} \end{pmatrix}=B^{\mathrm{T}}$$

（3）由矩阵乘积和矩阵迹的定义，得

$$\mathrm{tr}(X^{\mathrm{T}}AX)=(\sum_{i=1}^{m}a_{1i}x_{i1})x_{11}+(\sum_{i=1}^{m}a_{2i}x_{i1})x_{21}+\cdots$$

$$+(\sum_{i=1}^{m}a_{mi}x_{i1})x_{m1}+(\sum_{i=1}^{m}a_{1i}x_{i2})x_{12}+(\sum_{i=1}^{m}a_{2i}x_{i2})x_{22}+\cdots$$

$$+(\sum_{i=1}^{m}a_{mi}x_{i2})x_{m2}+\cdots+(\sum_{i=1}^{m}a_{1i}x_{in})x_{1n}+(\sum_{i=1}^{m}a_{2i}x_{in})x_{2n}+\cdots$$

$$+(\sum_{i=1}^{m}a_{mi}x_{in})x_{mn}=\sum_{h=1}^{n}\sum_{j=1}^{m}\sum_{i=1}^{m}a_{ji}x_{ih}x_{jh}$$

按定义3 即(3-21) 式，有

$$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(X^{\mathrm{T}}AX)$$

$$=\begin{pmatrix} \sum_{i=1}^{m}(a_{1i}+a_{i1})x_{i1} & \sum_{i=1}^{m}(a_{1i}+a_{i1})x_{i2} & \cdots \\ \sum_{i=1}^{m}(a_{2i}+a_{i2})x_{i1} & \sum_{i=1}^{m}(a_{2i}+a_{i2})x_{i2} & \cdots \\ \cdots\quad\cdots & \cdots\quad\cdots & \cdots \\ \sum_{i=1}^{m}(a_{mi}+a_{im})x_{i1} & \sum_{i=1}^{m}(a_{mi}+a_{im})x_{i2} & \cdots \end{pmatrix}$$

$$\begin{pmatrix} \sum_{i=1}^{m}(a_{1i}+a_{i1})x_{in} \\ \sum_{i=1}^{m}(a_{2i}+a_{i2})x_{in} \\ \cdots \qquad \cdots \\ \sum_{i=1}^{m}(a_{mi}+a_{im})x_{in} \end{pmatrix} = (A+A^{T})X$$

证毕。

此外，还可以验证纯量函数$f(X)$、$g(X)$对矩阵X有如下运算法则：

$$\frac{d}{dX}(f+g)=\frac{df}{dX}+\frac{dg}{dX} \tag{3-25}$$

$$\frac{d}{dX}(fg)=\frac{df}{dX}g+f\frac{dg}{dX} \tag{3-26}$$

注意：公式中乘积可以交换。

由于向量是矩阵的特殊情形，所以上述定义对向量也适用。因为对向量求导是常用的，故有必要讨论纯量函数对向量的导数。

定义4　设$f(X)$是以n维向量$X=[x_1, x_2, \cdots, x_n]^T$的$n$个分量$x_i$为自变量的可微纯量函数。即

$$f(X)=f(x_1, x_2, \cdots, x_n)$$

则
$$\frac{df}{dX}=\begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \tag{3-27}$$

称为纯量函数对n维向量X的导数，仍然是n维向量，记为

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\left(\frac{\partial f}{\partial x_i}\right)_{n\times 1} \qquad (i=1, 2, \cdots, n)$$

〔例3〕 设a是n维常向量，X为n维变向量，即

$$a=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \qquad X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

求线性型

$$f=a^{\mathrm{T}}X=a_1x_1+a_2x_2+\cdots+a_nx_n$$

对向量X的导数。

解 由定义4，有 $\frac{\partial f}{\partial x_i}=a_i$，则

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}=a$$

也即

$$\frac{\mathrm{d}}{\mathrm{d}X}(a,X)=\frac{\mathrm{d}}{\mathrm{d}X}a^{\mathrm{T}}X=a$$

由定义4可验证二次型的微分公式

$$\frac{\mathrm{d}}{\mathrm{d}x}X^{\mathrm{T}}AX=2AX \tag{3-28}$$

事实上，设

$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; \qquad A=[a_{ij}]_{n\times n}$$

且$a_{ij}=a_{ji}$ $(i\neq j)$ 由二次型

$$X^{T}AX=(\sum_{i=1}^{n}a_{1i}x_i)x_1+(\sum_{i=1}^{n}a_{2i}x_i)x_2+\cdots+\sum_{i=1}^{n}a_{ni}x_i)x_n$$

按(3-27)式，得

$$\frac{\mathrm{d}f}{\mathrm{d}X}=\begin{pmatrix}2a_{11}x_1+(a_{12}+a_{21})x_2+\cdots+(a_{1n}+a_{n1})x_n\\(a_{21}+a_{12})x_1+2a_{22}x_2+\cdots+(a_{2n}+a_{n2})x_n\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\(a_{n1}+a_{1n})x_1+(a_{n2}+a_{2n})x_2+\cdots+2a_{nn}x_n\end{pmatrix}$$

由$a_{ij}=a_{ji}(i\neq j)$，得

$$\frac{\mathrm{d}f}{\mathrm{d}X}=2\begin{pmatrix}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\a_{12}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\a_{1n}x_1+a_{2n}x_2+\cdots+a_{nn}x_n\end{pmatrix}=2AX$$

三、向量函数关于向量的微分

定义5 设m个多元函数

$$\begin{array}{l}z_1\ (x_1,\ x_2,\ \cdots,x_n)\\z_2\ (x_1,\ x_2,\ \cdots,\ x_n)\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\z_m\ (x_1,\ x_2,\ \cdots,\ x_n)\end{array}\qquad(3\text{-}29)$$

对自变量x_1，$x_2,\cdots,x_n$的偏导数都存在，记向量

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix};\qquad Z=\begin{pmatrix}z_1\\z_2\\\vdots\\z_m\end{pmatrix}$$

则（3-29）式可表为向量的形式：

$$Z(X)=\begin{pmatrix} z_1(X) \\ z_2(X) \\ \vdots \\ z_m(X) \end{pmatrix}$$

于是我们可以定义,

$$\frac{dZ}{dX}=\begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \frac{\partial z_1}{\partial x_2} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \frac{\partial z_2}{\partial x_1} & \frac{\partial z_2}{\partial x_2} & \cdots & \frac{\partial z_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial z_m}{\partial x_1} & \frac{\partial z_m}{\partial x_2} & \cdots & \frac{\partial z_m}{\partial x_n} \end{pmatrix} \tag{3-30}$$

称为m维向量$Z(X)$对n维向量X的导数。它是一个$m\times n$阶矩阵，且第i行等于$\left(\frac{dz_i}{dX}\right)^T$

〔例4〕 求线性向量函数$Y=AX$对向量X的导数。

其中

$$Y=\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix};\quad X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix};\quad A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

解 因为$y_i=\sum\limits_{j=1}^{n} a_{ij}x_j$, $\quad i=1, 2, \cdots, m$

于是

$$\frac{\mathrm{d}y_i}{\mathrm{d}X}=\begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}, \qquad i=1, 2, \cdots, m$$

则

$$\frac{\mathrm{d}Y}{\mathrm{d}X}=\begin{pmatrix} \left(\frac{\mathrm{d}y_1}{\mathrm{d}X}\right)^{\mathrm{T}} \\ \left(\frac{\mathrm{d}y_2}{\mathrm{d}X}\right)^{\mathrm{T}} \\ \vdots \\ \left(\frac{\mathrm{d}y_m}{\mathrm{d}X}\right)^{\mathrm{T}} \end{pmatrix}=A$$

关于n元函数$y=f(X)$对n维向量X的二阶导数由(3-27)及定义5中（3-30）式，得到

$$\frac{\mathrm{d}^2y}{\mathrm{d}X^2}=\frac{\mathrm{d}}{\mathrm{d}X}\left(\frac{\mathrm{d}y}{\mathrm{d}X}\right)=\begin{pmatrix} \frac{\partial^2 f}{\partial x_1{}^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n} \\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2{}^2} & \cdots & \frac{\partial^2 f}{\partial x_2\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \frac{\partial^2 f}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \tag{3-31}$$

对于任何可微的，具有连续混合二阶偏导数的多元函数$f(X)$而言，矩阵（3-31）是对称的。

我们研究多元函数的二阶Taylor展开式时，常用这样的矩阵来表达。

设$y=f(X)$

其中
$$X=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

由于
$$\frac{dy}{dX}=\begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$
对于多元函数$y=f(X)$在点
$$X_0=\begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{pmatrix}$$
的二阶Taylor展开式，可以写为：

$$f(X_0+\Delta X)=f(X_0)+\left(\frac{d}{dX}f(X)\Big|_{X=X_0}\right)^T\Delta X$$
$$+\frac{1}{2}(\Delta X)^T\frac{d_2}{dX^2}f(X)\Big|_{X=X_0}\Delta X+o[(\Delta X)^2] \qquad (3\text{-}32)$$

式中向量

$$\Delta X = X - X_0 = \begin{pmatrix} x_1 - x_1^0 \\ x_2 - x_2^0 \\ \vdots \\ x_n - x_n^0 \end{pmatrix}$$

表示多元函数自变量的增量向量。

对上述(3-32)式，可用来讨论多元函数极值的充分条件。因为多元函数$f(X)$在点X_0取得极值的必要条件是向量

$$\frac{\mathrm{d}}{\mathrm{d}X} f(X) \Big|_{X=X_0} = 0$$

即
$$\frac{\partial f}{\partial x_i} \Big|_{x_i = x_i^0} = 0 \quad (i=1,\ 2,\ \cdots,\ n)$$

至于在点X_0是否取得极值，是极大值或极小值，就可以从公式（3-32）来确定。由于（3-32）式中第二项

$\left(\frac{\mathrm{d}}{\mathrm{d}X} f(X) \Big|_{X=X_0}\right)^{\mathrm{T}} \Delta X = 0$，所以$f(X)$在点$X_0$取极大值或极小值取决于第三项。若对任意的$\Delta X \neq 0$，都有$\frac{1}{2}(\Delta X)^{\mathrm{T}} \times \frac{\mathrm{d}^2}{\mathrm{d}X^2} f(X) \Big|_{X=X_0} \Delta X > 0$，即$\frac{\mathrm{d}^2}{\mathrm{d}X^2} f(X) \Big|_{X=X_0}$为正定，则$\Delta f = f(+\Delta X) - f(X_0) > 0$，函数$f(X)$在点$X_0$必取得极小值。反之，若对任意的$\Delta X \neq 0$，都有$\frac{1}{2}(\Delta X)^{\mathrm{T}} \times \frac{\mathrm{d}^2}{\mathrm{d}X^2} f(X) \Big|_{X=X_0} \Delta X < 0$，即$\frac{\mathrm{d}^2}{\mathrm{d}X^2} f(X) \Big|_{X=X_0}$为负定，则$\Delta f < 0$，函数$f(X)$在点$X_0$必取得极大值，这就是函数$f(X)$在点$X_0$具有极值的充分

条件。

本节所述结论在解线性微分方程，矩阵微分方程和过程控制中都有重要应用。

第三节 向量和矩阵的范数

一、向量的范数（norm）

向量范数的概念是由推广R^1或R^2中的向量长度概念而产生的。范数最简单的例子是绝对值函数，它表示一个数到原点的距离。

例如，数ξ的绝对值$|\xi|$可用$|\xi|=\sqrt{\xi^2}$来定义；二维Euclid空间R^2上，向量$\eta=(\eta_1,\eta_2)$的欧氏长度$\|\eta\|$，其定义为：

$$\|\eta\|=\sqrt{\eta_1^2+\eta_2^2}$$

显然$\|\eta\|$是点η（η_1，η_2）到原点的距离，表示了向量η的大小。该量具有和实数及复数的模相似的性质。即具有：

（1）若$\eta\neq0$时，则$\|\eta\|>0$，当$\eta=0$时，则$\|\eta\|=0$；

（2）$\|a\eta\|=|a|\|\eta\|$，a为任意常数；

（3）对于两个向量η，ξ存在下列关系：

$$\|\eta+\xi\|\leqslant\|\eta\|+\|\xi\| \qquad \text{（三角不等式）}$$

对于一般的线性空间，如何来定义“长度”？这就引进了比长度更加广泛的概念——范数（norm）。

定义1 设V是数域F上的一个线性空间，$X\in V$是任意一个向量，X对应一个满足下列条件的非负实值函数$\|X\|$。

（1）当$X\neq0$时，$\|X\|>0$，当$X=0$时，$\|X\|=0$；

（2）$\|aX\|=|a|\,\|X\|$，其中$a\in F$；

（3）$\|X+Y\|\leqslant\|X\|+\|Y\|$，其中$X$，$Y\in V$

则称函数$\|X\|$为V中**向量X的范数**。

根据定义容易推出向量范数的一些基本性质：

性质1 零向量的范数是零。即 $\|0\|=0$

性质2 对任意$X\in V$，有 $\|-X\|=\|X\|$

性质3 当$X\neq 0$时，有 $\left\|\dfrac{X}{\|X\|}\right\|=1$

性质4 对任意X，$Y\in V$，有 $\big|\|X\|-\|Y\|\big|\leqslant\|X-Y\|$

证明 由定义1中的条件(3)，有

$$\|X\|=\|X-Y+Y\|\leqslant\|X-Y\|+\|Y\|$$

从而 $$\|X-Y\|\geqslant\|X\|-\|Y\|$$

又根据性质2，有

$$\|X-Y\|=\|Y-X\|\geqslant\|Y\|-\|X\|$$

则 $$-\|X-Y\|\leqslant\|X\|-\|Y\|$$

故 $$\big|\|X\|-\|Y\|\big|\leqslant\|X-Y\|$$

[**例1**] 证明R^3上，向量$(\xi=\xi_1,\ \xi_2,\ \xi_3)$的长度是$R^3$上向量$\xi$的一种范数。

证明 只须验证$\|\xi\|$满足范数的三个条件即可。

(1)和(2)显然成立。只验证(3)，即

对R^3中任意向量$\xi=(\xi_1,\ \xi_2,\ \xi_3)$，$\eta=(\eta_1,\ \eta_2,\ \eta_3)$ 要证明有

$$\|\xi+\eta\|\leqslant\|\xi\|+\|\eta\|\text{成立。}$$

因为 $\xi+\eta=(\xi_1+\eta_1,\ \xi_2+\eta_2,\ \xi_3+\eta_3)$

于是有 $$\|\xi+\eta\|=\sqrt{(\xi_1+\eta_1)^2+(\xi_2+\eta_2)^2+(\xi_3+\eta_3)^2}$$

而 $\|\xi\|=\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$，$\|\eta\|=\sqrt{\eta_1^2+\eta_2^2+\eta_3^2}$

由于对一切实数ξ_i和η_i（$i=1,\ 2,\ 3$）有

$$\left(\sqrt{(\xi_1+\eta_1)^2+(\xi_2+\eta_2)^2+(\xi_3+\eta_3)^2}\right)^2$$

$$=(\xi+\eta_1)^2{}_1+(\xi_2+\eta_2)^2+(\xi_3+\eta_3)^2$$

$$=\xi_1^2+\xi_2^2+\xi_3^2+2(\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3)+\eta_1^2+\eta_2^2+\eta_3^2$$

又R^3中内积

$$(\xi,\ \eta)=\xi_1\eta_1+\xi_2\eta_2+\xi_3\eta_3 \quad 及 \quad (\xi,\ \eta)\leqslant\|\xi\|\|\eta\|$$

得到

$$\|\xi+\eta\|^2=\|\xi\|^2+2(\xi,\eta)+\|\eta\|^2\leqslant\|\xi\|^2+2\|\xi\|\|\eta\|+\|\eta\|^2$$
$$=(\|\xi\|+\|\eta\|)^2$$

则 $\|\xi+\eta\|\leqslant\|\xi\|+\|\eta\|$

故 $\|\xi\|=\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$ 是R^3上向量ξ的一种范数。

[例2] 在R^3上还可以定义另一种范数，如对于向量$\xi=(\xi_1,\xi_2,\xi_3)$，令$\|\xi\|=\max\limits_i|\xi_i|$，则$\|\xi\|$满足定义1中的三个条件。

证明 (1) 当$\xi\neq0$时，有$\|\xi\|=\max\limits_i|\xi_i|\geqslant|\xi_i|>0$当$\xi=0$时，显然$\|\xi\|=0$；

(2) 对任意实数a，有

$a\xi=(a\xi_1,\ a\xi_2,\ a\xi_3)$，则

$\|a\xi\|=\max\limits_i|a\xi_i|=|a|\max\limits_i|\xi_i|=|a|\|\xi\|$

(3) 对任意向量$\xi=(\xi_1,\ \xi_2,\ \xi_3)$，$\eta=(\eta_1,\ \eta_2,\ \eta_3)$

有$\|\xi+\eta\|=\max\limits_i|\xi_i+\eta_i|\leqslant\max\limits_i|\xi_i|+\max\limits_i|\eta_i|$

则 $\|\xi+\eta\|\leqslant\|\xi\|+\|\eta\|$

故 $\|\xi\|=\max\limits_i\left\{|\xi_i|;\ i=1,\ 2,\ 3\right\}=\max\limits_i|\xi_i|$ 也是R^3上向量ξ的一种范数。

从上面二例可看到，一个向量空间可以定义多种范数。

现在我们讨论经常使用的n维空间R^n上的三种范数。

$$\text{I} \qquad \|\xi\|_1=\sum_{i=1}^{n}|\xi_i| \tag{3-33}$$

可证明$\|\xi\|_1$满足定义1中的三个条件。

(1) 当$\xi \neq 0$时，有$\|\xi\|_1 = \sum_{i=1}^{n} |\xi_i| > 0$

当$\xi = 0$时，显然$\|\xi\|_1 = 0$；

(2) 对任意$\alpha \in R$，有

$$\|\alpha\xi\|_1 = \sum_{i=1}^{n} |\alpha\xi_i| = |\alpha| \sum_{i=1}^{n} |\xi_i| = |\alpha| \, \|\xi\|_1;$$

(3) 对于任意向量ξ，$\eta \in R^n$，有

$$\|\xi+\eta\|_1 = \sum_{i=1}^{n} |\xi_i+\eta_i| \leqslant \sum_{i=1}^{n} (|\xi_i| + |\eta_i|)$$

$$= \sum_{i=1}^{n} |\xi_i| + \sum_{i=1}^{n} |\eta_i|$$

则 $\|\xi+\eta\|_1 \leqslant \|\xi\|_1 + \|\mu\|_1$

故 $\|\xi\|_1$是R^n上向量ξ的一种范数。称为**1-范数**

Ⅱ $\|\xi\|_2 = \sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2}$ (3-34)

相当于在R^n上取内积为$(\xi, \eta) = \sum_{i=1}^{n} \xi_i \eta_i$时的长度。和例1证法一样，$\|\xi\|_2$满足定义1中的三个条件。

故 $\|\xi\|_2$也是R^n上向量ξ的一种范数。称为**2-范数**。

Ⅲ $\|\xi\|_\infty = \max_i |\xi_i|$ (3-35)

也和例2证法一样，$\|\xi\|_\infty$满足定义1中的三个条件。

故 $\|\xi\|_\infty$也是R^n上向量ξ的一种范数。称为**∞-范数**。

此三种范数$\|\xi\|_1$，$\|\xi\|_2$，$\|\xi\|_\infty$都可由下列范数表示

$$\|\xi\|_p = \left(\sum_{i=1}^{n} |\xi_i|^p\right)^{\frac{1}{p}} \quad 1 \leqslant p < \infty \qquad (3\text{-}36)$$

对任意实数$p \geqslant 1$，上式都满足定义1中的三个条件（读者自证）。

对于R^n中这样的范数简称为**p-范数**

当$p=1$时，$\|\xi\|_p=\|\xi\|_1$；

当$p=2$时，$\|\xi\|_p=\|\xi\|_2$；

当$p\to\infty$时，有$\|\xi\|_p\to\|\xi\|_\infty$，即$\lim\limits_{p\to\infty}\|\xi\|_p=\|\xi\|_\infty$

对上式证明如下：

$$\|\xi\|_p=\left(\sum_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}},\qquad \xi\neq 0$$

若 $|\xi_{i0}|=\max\{|\xi_1|,|\xi_2|,\cdots,|\xi_n|\}$，而$|\xi_{i0}|\neq 0$，

则 $\|\xi\|_\infty=\max\limits_i|\xi_i|=|\xi_{i0}|$

又 $\|\xi\|_p=\left(\sum\limits_{i=1}^{n}|\xi_{i0}|^p\dfrac{|\xi_i|^p}{|\xi_{i0}|^p}\right)^{\frac{1}{p}}=|\xi_{i0}|\left(\sum\limits_{i=1}^{n}\dfrac{|\xi_i|^p}{|\xi_{i0}|^p}\right)^{\frac{1}{p}}$

由于 $|\xi_{i0}|^p\leqslant\sum\limits_{i=1}^{n}|\xi_i|^p\leqslant n|\xi_{i0}|^p$，

所以 $1\leqslant\dfrac{1}{|\xi_{i0}|}\left(\sum\limits_{i=1}^{n}|\xi_i|^p\right)^{\frac{1}{p}}\leqslant n^{\frac{1}{p}}$，

而 $\lim\limits_{p\to\infty}n^{\frac{1}{p}}=1$，

故 $\lim\limits_{p\to\infty}\|\xi\|_p=\|\xi\|_\infty$

此外，在其它线性空间上也可以定义各种各样的范数。

例如，在有限闭区间$[t_0, t_1]$上定义的连续实函数的全体构成R上的向量空间。由上述类推，可定义

$$\|f\|_1=\int_{t_0}^{t_1}|f(t)|dt$$

$$\|f\|_p=\left[\int_{t_0}^{t_1}|f(t)|^p dt\right]^{\frac{1}{p}},\qquad 1\leqslant p<\infty \tag{3-37}$$

$$\|f\|_\infty=\max_{t\in[t_0,t_1]}|f|$$

对于$\|f\|_1$，$\|f\|_p$，$\|f\|_\infty$都满足定义1中的三个条件（读者自证）。

由于在有限维线性空间上范数的种数有无穷多，而这些范数之间有重要的关系。

定理1 有限维线性空间上的不同范数是等价的。即是：对任何有限维线性空间V上所定义的任意两种范数$\|X\|_a$，$\|X\|_b$都存在两个与X无关的正常数c_1和c_2，使

$$\|X\|_a \leqslant c_1\|X\|_b, \qquad \|X\|_b \leqslant c_2\|X\|_a \tag{3-38}$$

（对所有$X\in V$满足（3-38）式的两种范数称为是等价的）。

证明 设V是n维线性空间，e_1，e_2，…，e_n是V的一组基底。于是，对任意向量$X\in V$，有

$$X=x_1e_1+x_2e_2+\cdots+x_ne_n,$$

定义

$$\|X\|_E=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$$

显然它满足定义1中的三个条件，所以它是一种范数。又因

$$\|X\|_a=\|x_1e_1+x_2e_2+\cdots+x_ne_n\|_a$$

可定义 $\|X\|_a=\varphi\ (x_1,\ x_2\cdots x_n)$

可以证明$\varphi(x_1,\ x_2\cdots,\ x_n)$是$(x_1,\ x_2,\ \cdots,\ x_n)$的连续函数。

为此，设另一个向量$X'\in V$，有

$$X'=x_1'e_1+x_2'e_2+\cdots+x_n'e_n$$

于是 $\|X'\|_a=\varphi\ (x_1',\ x_2',\ \cdots,x_n')$

则 $|\varphi\ (x_1',\ x_2',\ \cdots,\ x_n')-\varphi(x_1,\ x_2,\ \cdots,\ x_n)|=|\,\|X'\|_a-\|X\|_a\,|\leqslant\|X'-X\|_a=\|(x_1'-x_1)\ e_1+\ (x_2'-x_2)e_2+\cdots+\ (x_n'-x_n)e_n\|_a\leqslant|x_1'-x_1|\|e_1\|_a+|x_2'-x_2|\|e_2\|_a+\cdots+|x_n'-x_n|\|e_n\|_a$ 因$\|e_i\|_a$ $(i=1,2,\cdots n,)$是常数，所以当x_i与X_i'充分接近时，$\varphi(x_1',\ x_2',\ \cdots,\ x_n')$就充分接近$\varphi(x_1,\ x_2,$

$\cdots, x_n)$。故$\varphi(x_1, x_2, \cdots, x_n)$是连续函数。

再由连续函数的性质，可知函数$\varphi(x_1, x_2, \cdots, x_n)$在有界闭集

$$x_1^2+x_2^2+\cdots+x_n^2=1 \tag{3-39}$$

上（即Euclid空间R^n的球面上），存在着最大值M和最小值m。

又因（3-39）式中的x_i不能全为零，所以$m>0$，设

$$d=\sqrt{\sum_{i=1}^{n} x_i^2} \quad \text{及} \quad Y=\frac{1}{d}X, \text{ 即}$$

$$Y=\frac{x_1}{d}e_1+\frac{x_2}{d}e_2+\cdots+\frac{x_n}{d}e_n$$

则向量$Y\in V$。而

$$\left(\frac{x_1}{d}\right)^2+\left(\frac{x_2}{d}\right)^2+\cdots+\left(\frac{x_n}{d}\right)^2=1$$

于是，在球面（3-39）上，有

$$0<m\leqslant\|Y\|_a=\varphi\left(\frac{x_1}{d}, \frac{x_2}{d}, \cdots, \frac{x_n}{d}\right)\leqslant M$$

又因$Y=\frac{1}{d}X$，则有

$$md\leqslant\|X\|_a\leqslant Md, \quad \text{即}$$

$$m\|X\|_E\leqslant\|X\|_a\leqslant M\|X\|_E$$

取$c_1'=M$，$c_2'=\frac{1}{m}$，则有

$$\|X\|_a\leqslant c_1'\|X\|_E, \quad \|X\|_E\leqslant c_2'\|X\|_a$$

类似可得

$$\|X\|_b\leqslant d_1\|X\|_E, \quad \|X\|_E\leqslant d_2\|X\|_b$$

于是有

$$\|X\|_a\leqslant c_1'\|X\|_E\leqslant c_1' d_2\|X\|_b$$

$\|X\|_b \leqslant d_1\|X\|_c \leqslant d_1 c_2'\|X\|_a$

取　$c_1 = c_1' d_2$，$c_2 = d_1 c_2'$，即得（3-38）式。

由此不难看出，范数的等价关系具有传递性。

定理2　R^n中向量$X^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$收敛于$X = (x_1, x_2, \cdots, x_n)$的充要条件是：对任意一种范数$\|\cdot\|$，均有序列$\left\{\|X^{(k)} - X\|\right\}$收敛于零。

即　$\lim\limits_{k\to\infty}\|X^{(k)} - X\| = 0$

证明　利用范数的等价性，只需对一种范数作证明，则任意一种范数都成立。不妨取$\|\cdot\| = \|\cdot\|_\infty$

由向量序列$\{X^{(k)}\}$收敛定义，有

$$\lim_{k\to\infty} X^{(k)} = X \Longleftrightarrow \lim_{k\to\infty} x_i^{(k)} = x_i \quad (i = 1, 2, \cdots, n)$$

$$\Longleftrightarrow \lim_{k\to\infty} |x_i^{(k)} - x_i| = 0 \Longleftrightarrow \lim_{k\to\infty} \max_{1\leqslant i\leqslant n} |x_i^{(k)} - x_i| = 0$$

$$\Longleftrightarrow \lim_{k\to\infty} \|X^{(k)} - X\| = 0$$

故序列$\{\|X^{(k)} - X\|\}$收敛于零是向量$X^{(k)}$收敛于X的充要条件。

二、矩阵的范数

我们可以把向量范数的概念扩充到矩阵上，因为所有$m\times n$阶矩阵构成一个线性空间，在本质上和mn维的向量空间是恒等的。于是当把一个$m\times n$阶矩阵视为mn维的向量时，当然也可以定义范数。但是为了考虑矩阵的乘积，令$m = n$

定义2　在数域F上，n阶矩阵的全体所构成的线性空间$F^{n\times n}$中，规定一个非负实值函数$\|A\|$，如果对任意A，$B \in F^{n\times n}$满足下列四个条件：

（1）若$A \neq 0$，则$\|A\| > 0$，当$A = 0$时，则$\|0\| = 0$；

（2）对任意$a \in F$，则$\|aA\| = |a|\ \|A\|$；

(3) $\|A+B\| \leqslant \|A\|+\|B\|$;

(4) $\|AB\| \leqslant \|A\|\|B\|$

则此非负实值函数$\|A\|$就称为n阶**矩阵A的范数**。

由于在$R^{n\times n}$上的矩阵范数，本质上是R^{n^2}上的向量范数。从而可以和向量一样证明$R^{n\times n}$上的n阶矩阵序列$\{A^{(k)}\} \longrightarrow A$ $(k\rightarrow\infty)$的充分必要条件是：

$$\|A^{(k)}-A\| \longrightarrow 0 \quad (k\rightarrow\infty)$$

由定义2中(3)，可得 $\|A-B\| \geqslant \left| \|A\|-\|B\| \right|$

这样就说明了矩阵范数是具有连续性的。即可由$A^{(k)} \longrightarrow A$ $(k\rightarrow\infty)$推出$\|A^{(k)}\| \rightarrow \|A\|$ $(k\rightarrow\infty)$

同向量一样，矩阵也可以有各种各样的范数，一般常把矩阵的范数和向量的范数混在一起，故有必要考虑矩阵和向量的积。为此，引进范数相容的概念。

定义3 如果任意向量$X\in R^n$，和任意n阶矩阵$A\in R^{n\times n}$，对给定的向量范数$\|X\|$和矩阵范数$\|A\|$满足不等式

$$\|AX\| \leqslant \|A\|\|X\| \qquad (3\text{-}40)$$

则称矩阵范数$\|A\|$和向量范数$\|X\|$**相容** (Compatible)。

一般地，当向量$X\in R^n$的范数$\|X\|$和矩阵$A\in R^{n\times n}$给定时，必然存在和$\|X\|$相容的矩阵A的范数$\|A\|$。

例如，可用下式来定义这种矩阵范数：

$$\|A\| = \max_{\|X\|=1} \|AX\| \qquad (3\text{-}41)$$

其中向量X取遍R^n中范数为1的所有向量的集合。

根据向量范数是其分量的连续函数这个性质，则对每一个阵矩A来讲，(3-41)式中的最大值都是可以达到的。即是说：总可以找到这样的向量$X_0\neq 0$，$\|X_0\|=1$，使$\|AX_0\| = \|A\|$。

下面证明，按这样定义的矩阵范数，满足定义 2 中的四个条件及定义3中的相容条件。

(1) 设$A\neq 0$，则一定存在着$\|X\|=1$ 的向量X，使得$AX\neq 0$，

因此 $\|AX\|>0$，

所以 $\|A\|=\max\limits_{\|X\|=1}\|AX\|>0$

若$A=0$，则必定有$\|A\|=\max\limits_{\|X\|=1}\|0X\|=0$

(2) 由(3-41)式，知

$$\|\alpha A\|=\max_{\|X\|=1}\|\alpha AX\|=|\alpha|\max_{\|X\|=1}\|AX\|=|\alpha|\,\|A\| \quad (\alpha\in F)$$

验证相容条件：

设$Y\neq 0$为任意一个向量，则$X=\dfrac{1}{\|Y\|}Y$满足

$$\|X\|=1$$

于是 $\|AY\|=\|A(\|Y\|X)\|=\|Y\|\|AX\|\leqslant\|A\|\|Y\|$

这就是相容条件。

(3) 对矩阵$A+B$，总存在着向量X_0，且$\|X_0\|=1$，使得

$$\|A+B\|=\|(A+B)X_0\|$$

于是 $\|A+B\|=\|(A+B)X_0\|=\|AX_0+BX_0\|$

$\leqslant\|AX_0\|+\|BX_0\|\leqslant\|A\|\|X_0\|+\|B\|\|X_0\|$

故 $\|A+B\|\leqslant\|A\|+\|B\|$

(4) 对矩阵AB，总存在着向量X_0，且$\|X_0\|=1$，

使得 $\|AB\|=\|ABX_0\|$

于是 $\|AB\|=\|A(BX_0)\|\leqslant\|A\|\|BX_0\|\leqslant\|A\|\|B\|\|X_0\|$

故 $\|AB\|\leqslant\|A\|\|B\|$ 证毕。

因此按 (3-41)式定义的矩阵范数，我们称为由向量范数$\|X\|$导出的范数 (induced norm)。

实际上，当把矩阵看成线性算子时，这样的范数有时又称为从属于向量范数$\|X\|$的**算子范数**。特别是对于$A=I$，显然$\|I\|=1$。

现在我们利用前面所述的三种常用的向量范数导出**矩阵范数**（从属范数或**导出范数**）。

定理3　设$A=[a_{ij}]_{n\times n}$，则

Ⅰ、$\|X\|_1=\sum_{i=1}^{n}|x_i|$导出的矩阵范数是：

$$\|A\|_1=\max_{\|X\|_1=1}\|AX\|_1=\max_{j}\sum_{i=1}^{n}|a_{ij}| \tag{3-42}$$

称为矩阵的1-范数（即矩阵的列向量的1-范数的最大值）。

Ⅱ、$\|X\|_2^2=\sum_{i=1}^{n}|x_i|^2=(X,\ X)$导出的矩阵范数是：

$$\|A\|_2=\max_{\|X\|_2=1}\|AX\|_2=\sqrt{\lambda_1} \tag{3-43}$$

其中λ_1是矩阵A^TA的最大特征值。

称$\|A\|_2$为矩阵的2-范数。

Ⅲ、$\|X\|_\infty=\max_{i}|x_i|$导出的矩阵范数是：

$$\|A\|_\infty=\max_{\|X\|_\infty=1}\|AX\|_\infty=\max_{i}\sum_{j=1}^{n}|a_{ij}| \tag{3-44}$$

称为矩阵的∞-范数（即矩阵的行向量的1-范数的最大值）。

证明　Ⅰ、设$\|X\|_1=1$，于是

$$\|AX\|_1=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right|\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}||x_j|$$

$$=\sum_{j=1}^{n}|x_j|\left(\sum_{i=1}^{n}|a_{ij}|\right)\leqslant\left(\mathrm{mxa}_{j}\sum_{i=1}^{n}|a_{ij}|\right)\sum_{j=1}^{n}|x_j|$$

$$=\max_j \sum_{i=1}^{n} |a_{ij}| \|X\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

另一方面，若$j=k$时，$\sum_{i=1}^{n}|a_{ij}|$达到最大值。令

$X_0=(x_1^{(0)},\ x_2^{(0)},\ \cdots,\ x_n^{(0)})$,

其中$x_j^{(0)}=\begin{cases}0 & j\neq k\\ 1 & j=k\end{cases}\quad (j=1,\ 2,\ \cdots,\ n)$

显然对这样作的向量X_0，有$\|X_0\|_1=1$，从而

$$\|AX_0\|_1=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}X_j^{(0)}\right|=\sum_{i=1}^{n}\left|a_{ik}\right|=\max_j\sum_{i=1}^{n}\left|a_{ij}\right|$$

于是 $\max\limits_{\|X\|_1=1}\|AX\|_1=\max\limits_j\sum_{i=1}^{n}|a_{ij}|$

故 $\|A\|_1=\max\limits_j\sum_{i=1}^{n}|a_{ij}|$

Ⅱ、由(3-41)式，有$\|A\|_2=\max\limits_{\|X\|_2=1}\|AX\|_2$

由于A^TA是对称矩阵，其对应的二次型

$f(X)=X^TA^TAX=(AX)^T(AX)\geqslant 0$

因此，A^TA的n个特征值每个都大于或等于零。

设$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n\geqslant 0$是$A^TA$的$n$个特征值，而$X_1,X_2,\cdots,X_n$是对应于它们的一组标准正交特征向量。则对任何一个具有$\|X\|_2=1$的向量X，都可表示为：

$X=a_1X_1+a_2X_2+\cdots+a_nX_n$

于是 $(X,\ X)=a_1^2+a_2^2+\cdots+a_n^2=1$

又 $\|AX\|_2^2=(AX,\ AX)=(X,\ A^TAX)$

$=(a_1X_1+a_2X_2+\cdots+a_nX_n)^T(\lambda_1a_1X_1+\lambda_2a_2X_2+\cdots+\lambda_na_nX_n)$

$=\lambda_1a_1^2+\lambda_2a_2^2+\cdots+\lambda_na_n^2\leqslant\lambda_1(a_1^2+a_2^2+\cdots+a_n^2)$

$=\lambda_1$

而对向量$X=X_1$，有

$$\|AX_1\|_2^2=(AX_1, AX_1)=(X_1, A^TAX_1)=(X_1, \lambda_1X_1)$$
$$=\lambda_1$$

故 $\|A\|_2=\max\limits_{\|X\|_2=1}\|AX\|_2=\sqrt{\lambda_1}$

Ⅲ、设$\|X\|_\infty=1$，

于是

$$\|AX\|_\infty=\max_i\left|\sum_{j=1}^n a_{ij}x_j\right|\leqslant\max_i\sum_{j=1}^n|a_{ij}||x_j|$$

$$\leqslant\max_i\sum_{j=1}^n|a_{ij}|\max_j|x_j|=\max_i\sum_{j=1}^n|a_{ij}|\|X\|_\infty$$

$$=\max_i\sum_{j=1}^n|a_{ij}|$$

因此 $\max\limits_{\|X\|_\infty=1}\|AX\|_\infty\leqslant\max\limits_i\sum\limits_{j=1}^n|a_{ij}|$

为了证明 $\max\limits_{\|X\|_\infty=1}\|AX\|_\infty=\max\limits_i\sum\limits_{j=1}^n|a_{ij}|$

只要作向量X_0，使得

$\|X_0\|_\infty=1$ 而

$\|AX_0\|_\infty=\max\limits_i\sum\limits_{j=1}^n|a_{ij}|$ 即可。

为此，设当$i=k$时，$\sum\limits_{j=1}^n|a_{ij}|$达到最大值，即

$$\max_i\sum_{j=1}^n|a_{ij}|=\sum_{j=1}^n|a_{kj}|$$

取向量X_0的第j个分量$x_j^{(0)}$如下：

$$x_j^{(0)}=\begin{cases}\dfrac{|a_{kj}|}{a_{kj}} & 当\quad a_{kj}\neq 0时\\ 1 & 当\quad a_{kj}=0时\end{cases}$$

显然有$\|X_0\|_\infty=1$，

$$当 i\neq k时\ \left|\sum_{j=1}^{n}a_{ij}x_j^{(0)}\right|\leqslant\sum_{j=1}^{n}|a_{ij}|\,|x_j^{(0)}|=\sum_{j=1}^{n}|a_{ij}|$$

$$\leqslant\sum_{j=1}^{n}|a_{kj}|$$

$$而\quad\left|\sum_{j=1}^{n}a_{kj}x_j^{(0)}\right|=\sum_{j=1}^{n}|a_{kj}|$$

$$于是\quad \max\left|\sum_{j=1}^{n}a_{ij}x_j^{(0)}\right|=\sum_{j=1}^{n}|a_{kj}|=\max_i\sum_{j=1}^{n}|a_{ij}|$$

$$则\quad \|AX_0\|_\infty=\max_i\left|\sum_{j=1}^{n}a_{ij}x_j^{(0)}\right|=\max_i\sum_{j=1}^{n}|a_{ij}|$$

$$故\quad \|A\|_\infty=\max_{\|X\|_\infty=1}\|AX\|_\infty=\max_i\sum_{j=1}^{n}|a_{ij}|\quad 证毕。$$

此外，若把矩阵A看着向量空间$F^{n\times n}$中的向量时，由向量的2-范数，可以引进一种常用的矩阵范数。

定义4 设矩阵$A=[a_{ij}]_{n\times n}$，$a_{ij}\in R$（或C），则定义

$$\|A\|_F=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}\qquad(3\text{-}45)$$

称为Frobenius范数，记作$\|\cdot\|_F$。显然满足

(1) 若$A\neq 0$，则$\|A\|_F>0$；

(2) $\|\alpha A\|_F=|\alpha|\,\|A\|_F$，$\alpha\in R$（或$C$）；

(3) $\|A+B\|_F\leqslant\|A\|_F+\|B\|_F$，

现在要证明

(4) $\|AB\|_F\leqslant\|A\|_F\|B\|_F$

对任意$A\in R^{n\times n}$（或$C^{n\times n}$）和$X\in R^n$，设A是按行分块的，即

$$A^T=[a_1,\ a_2,\ \cdots,\ a_n],\text{则}$$

$$AX=\begin{pmatrix}(a_1^T,\ X)\\(a_2^T,\ X)\\\vdots\\(a_n^T,\ X)\end{pmatrix}$$

其中$X=(x_1,\ x_2,\ \cdots,\ x_n)^T$；

$$(a_i^T,\ X)=a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n$$

由内积的性质，有

$$|(a_i^T,\ X)|\leqslant\|a_i\|_2\|X\|_2\ \text{(Cauchy—Schwarz 不等式)}$$

因此 $$\|AX\|_2^2=\sum_{i=1}^{n}|(a_i^T,\ X)|^2\leqslant\left(\sum_{i=1}^{n}\|a_i\|_2^2\right)\|X\|_2^2$$
$$=\|A\|_F^2\|X\|_2^2$$

则 $$\|AX\|_2\leqslant\|A\|_F\|X\|_2 \tag{3-46}$$

又设 $B\in F^{n\times n}$，且B是按列分块，即$B=[b_1,\ b_2,\cdots,b_n]$

则有$$\|AB\|_F^2=\|[Ab_1,\ Ab_2,\ \cdots,\ Ab_n]\|_F^2=\sum_{j=1}^{n}\|Ab_j\|_2^2$$

$$\leqslant\sum_{j=1}^{n}\|A\|_F^2\|b_j\|_2^2=\|A\|_F^2\sum_{j=1}^{n}\|b_j\|_2^2=\|A\|_F^2\|B\|_F^2$$

故 $$\|AB\|_F\leqslant\|A\|_F\|B\|_F$$

由(3-46)式，可见阵矩范数$\|A\|_F$与向量范数$\|\cdot\|_2$是相容的。

第四节 矩阵函数及其性质

由多项式 $f(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_k\lambda^k$， $a_i\in C$

$(i=1,2,\cdots,k)$，可定义n阶矩阵A的多项式

$$f(A)=a_0I+a_1A+a_2A^2+\cdots+a_kA^k$$

它是矩阵A的函数，称为矩阵函数。

定义矩阵函数的途径有好几种。例如，对任意$\lambda\in C$绝对收敛的纯量幂级数

$$e^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!},\quad \cos\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^{2k}}{(2k)!},$$

$$\sin\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^{2k+1}}{(2k+1)!}\text{及}$$

$$\ln(1+\lambda)=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^{k+1}}{k+1}\qquad(|\lambda|<1)$$

相应地，分别定义矩阵函数e^A，$\cos A$，$\sin A$及$\ln(I+A)$为下列绝对收敛的矩阵幂级数之和

$$e^A=\sum_{k=0}^{\infty}\frac{A^k}{k!}\qquad \cos A=\sum_{k=0}^{\infty}(-1)^k\frac{A^{2k}}{(2k)!}$$

$$\sin A=\sum_{k=0}^{\infty}(-1)^k\frac{A^{2k+1}}{(2k+1)!}\quad\text{及}$$

$$\ln(I+A)=\sum_{k=0}^{\infty}(-1)^k\frac{A^{k+1}}{k+1}$$

下面我们将从更一般的情况来定义矩阵函数，并由此定义可以得到矩阵函数的各种性质。矩阵函数的应用很广泛，下章将讨论矩阵函数在解微分方程组中的应用。这是现代控制理论的基础。

一、矩阵函数的概念

定义1 设n阶矩阵A的最小多项式为

$$\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_t)^{m_t}$$

其中λ_1，λ_2，…，λ_t是A的相异特征值。$f(\lambda)$为已知的任意解

析函数，若存在多项式$g(\lambda)$，使得$g(\lambda)$和$f(\lambda)$在矩阵A上的谱值相同（或者是在A的谱上一致），即有

$$f^{(l)}(\lambda_j)=g^{(l)}(\lambda_j) \qquad \begin{array}{l} j=1,\ 2,\ \cdots,\ t; \\ l=0,\ 1,\ 2,\ \cdots,\ m_j-1 \end{array}$$

则定义 $f(A)=g(A)$ (3-47)

称$f(A)$**为矩阵函数**。

显然，这样定义的矩阵函数，当$f(\lambda)$是多项式时，就和上章第三节中的定理5一致。

根据定义1，可以求出矩阵函数$f(A)$，并且可用$f(A)$来计算A^m，令$f(A)=A^m$即可。

［例1］ 设

$$A=\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}_{n\times n}$$

$f(\lambda)$为任意解析函数，求$f(A)$。

解 因为矩阵A的最小多项式为

$$\varphi_m(\lambda)=\frac{\psi(\lambda)}{D_{n-1}(\lambda)}=(\lambda-\lambda_i)^n$$

从而$f(\lambda)$在矩阵A上的谱值为$f(\lambda_i)$，$f'(\lambda_i)$，$\cdots$，$f^{(n-1)}(\lambda_i)$

按定义1要找一个多项式$g(\lambda)$，使得

$$g^{(k)}(\lambda_i)=f^{(k)}(\lambda_i),\quad k=0,\ 1,\ 2,\ \cdots,\ n-1$$

即是$g(\lambda)$与$f(\lambda)$在矩阵A上的谱值相同。因此$g(\lambda)$必定是λ的$n-1$次多项式。

设 $g(\lambda)=a_0+a_1(\lambda-\lambda_i)+a_2(\lambda-\lambda_i)^2+\cdots+a_{n-1}(\lambda-\lambda_i)^{n-1}$

由 $g(\lambda_i)=f(\lambda_i)$ 得 $a_0=f(\lambda_i)$

$$g'(\lambda_i)=f'(\lambda_i) \quad 得 \quad a_1=\frac{f'(\lambda_i)}{1!}$$

$$g''(\lambda_i)=f''(\lambda_i) \quad 得 \quad a_2=\frac{f''(\lambda_i)}{2!}$$

…………………………………

$$g^{(n-1)}(\lambda_i)=f^{(n-1)}(\lambda_i) \quad 得 \quad a_{n-1}=\frac{f^{(n-1)}(\lambda_i)}{(n-1)!}$$

则 $$g(\lambda)=f(\lambda_i)+f'(\lambda_i)(\lambda-\lambda_i)+\frac{f''(\lambda_i)}{2!}(\lambda-\lambda_i)^2+\cdots$$

$$+\frac{f^{(n-1)}(\lambda_i)}{(n-1)!}(\lambda-\lambda_i)^{n-1}$$

于是 $$f(A)=f(\lambda_i)I+f'(\lambda_i)(A-\lambda_iI)+\frac{f''(\lambda_i)}{2!}(A-\lambda_iI)^2$$

$$+\cdots+\frac{f^{(n-1)}(\lambda_i)}{(n-1)!}(A-\lambda_iI)^{n-1}$$

其中

$$A-\lambda_iI=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

故A的矩阵函数

$$f(A)=\begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \frac{1}{2!}f''(\lambda_i) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_i) \\ 0 & f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_i) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & f'(\lambda_i) \\ 0 & 0 & 0 & \cdots & f(\lambda_i) \end{pmatrix}$$

[例2] 设

$$A=\begin{pmatrix} 1 & -1 \\ -4 & 1 \end{pmatrix}$$

计算A^{20}

解 设$f(\lambda)=\lambda^{20}$，求$f(A)=A^{20}$

因为A的特征多项式

$$\psi(\lambda)=|\lambda I-A|=\begin{vmatrix} \lambda-1 & 1 \\ 4 & \lambda-1 \end{vmatrix}=(\lambda-3)(\lambda+1)$$

又 $D_1(\lambda)=1$，则A的最小多项式为

$$\varphi_m(\lambda)=\frac{\psi(\lambda)}{D_1(\lambda)}=(\lambda-3)(\lambda+1)$$

从而，令$g(\lambda)=a_0+a_1\lambda$

由于A的特征值$\lambda_1=3$，$\lambda_2=-1$，

于是由定义1，有

$g(3)=f(3)$，　　得　$a_0+3a_1=3^{20}$

$g(-1)=f(-1)$，　得　$a_0-a_1=(-1)^{20}$

解上方程组，得

$$a_1=\frac{3^{20}-1}{4}, \quad a_0=\frac{3^{20}+3}{4}$$

则 $g(\lambda)=\frac{3^{20}+3}{4}+\frac{3^{20}-1}{4}\lambda,$

故 $f(A)=A^{20}=g(A)=\frac{3^{20}+3}{4}I+\frac{3^{20}-1}{4}A$

$$=\begin{pmatrix}\frac{3^{20}+3}{4} & 0\\ 0 & \frac{3^{20}+3}{4}\end{pmatrix}+\frac{3^{20}-1}{4}\begin{pmatrix}1 & -1\\ -4 & 1\end{pmatrix}=\begin{pmatrix}\frac{3^{20}+1}{2} & \frac{1-3^{20}}{4}\\ 1-3^{20} & \frac{3^{20}+1}{2}\end{pmatrix}$$

[例3] 已知 $A=\begin{pmatrix}1 & -1\\ -4 & 1\end{pmatrix}$

证明：$\sin^2 A+\cos^2 A=I$

解 设 $f(\lambda)=\sin^2\lambda+\cos^2\lambda,$

由例2知A的最小多项式为：

$\varphi_m(\lambda)=(\lambda-3)(\lambda+1)$，$A$的特征值为$\lambda_1=3$，$\lambda_2=-1$，

于是，得到 $f(3)=\sin^2 3+\cos^2 3=1$

$f(-1)=\sin^2(-1)+\cos^2(-1)=1$

令 $g(\lambda)=a_0+a_1\lambda,$

由 $g(3)=f(3)$，得$a_0+3a_1=1$

$g(-1)=f(-1)$ 得 $a_0-a_1=1$

解上方程组，得 $a_0=1$，$a_1=0$

则 $g(\lambda)=1,$

故 $f(A)=\sin^2 A+\cos^2 A=g(A)=I$

此公式与普通三角公式一致。

同理可证 $e^{A}e^{-A}=I$

利用矩阵函数定义，还可以计算下列函数。

[例4] 设

$$A=\begin{pmatrix}2&1&4\\0&2&0\\0&3&1\end{pmatrix}$$

计算e^{At}和$\cos At$

解 因为A的特征多项式为

$$\psi(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda-2&-1&-4\\0&\lambda-2&0\\0&-3&\lambda-1\end{vmatrix}=(\lambda-2)^2(\lambda-1)$$

而 $D_2(\lambda)=1$，则A的最小多项式

$$\varphi_m(\lambda)=\frac{\psi(\lambda)}{D_2(\lambda)}=(\lambda-2)^2(\lambda-1)$$

设$f(\lambda)=e^{\lambda t}$，

取$g(\lambda)=a_0(t)+a_1(t)\lambda+a_2(t)\lambda^2$，

由定义1，得方程组：

$$a_0(t)+2a_1(t)+4a_2(t)=e^{2t}$$

$$a_1(t)+4a_2(t)=te^{2t}$$

$$a_0(t)+a_1(t)+a_2(t)=e^t$$

解之，得 $a_0(t)=4e^t-3e^{2t}+2te^{2t}$

$$a_1(t)=-4e^t+4e^{2t}-3te^{2t}$$

$$a_2(t)=e^t-e^{2t}+te^{2t}$$

则 $g(\lambda)=(4e^t-3e^{2t}+2te^{2t})+(-4e^t+4e^{2t}-3te^{2t})\lambda+(e^t-e^{2t}+te^{2t})\lambda^2$，

故 $f(A)=e^{At}=g(A)=(4e^t-3e^{2t}+2te^{2t})I+(-4e^t+4e^{2t}-3te^{2t})A+(e^t-e^{2t}+te^{2t})A^2$

$$=\begin{pmatrix}e^{2t}&12e^t-12e^{2t}+13te^{2t}&-4e^t+4e^{2t}\\0&e^{2t}&0\\0&-3e^t+3e^{2t}&e^t\end{pmatrix}$$

再设　$f(\lambda)=\cos\lambda t$，由定义1，得方程组

$$a_0(t)+2a_1(t)+4a_2(t)=\cos 2t$$

$$a_1(t)+4a_2(t)=-t\sin 2t$$

$$a_0(t)+a_1(t)+a_2(t)=\cos t$$

解之，得

$$a_0(t)=4\cos t-3\cos 2t-2t\sin 2t$$

$$a_1(t)=-4\cos t+4\cos 2t+3t\sin 2t$$

$$a_2(t)=\cos t-\cos 2t-t\sin 2t$$

故　$f(A)=\cos At=g(A)=(4\cos t-3\cos 2t-2t\sin 2t)I$

$+(-4\cos t+\cos 2t+3t\sin 2t)A+(\cos t-\cos 2t-t\sin 2t)A^2$

$$=\begin{pmatrix} \cos 2t & 12\cos t-12\cos 2t-13t\sin 2t & -4\cos t+4\cos 2t \\ 0 & \cos 2t & 0 \\ 0 & -3\cos t+3\cos 2t & \cos t \end{pmatrix}$$

二、矩阵函数的性质

性质1　A和$f(A)$可交换。即

$$Af(A)=f(A)A \qquad (3\text{-}48)$$

性质2　函数和(积)的矩阵函数等于矩阵函数的和(积)。即

$$(f_1+f_2)(A)=f_1(A)+f_2(A) \qquad (3\text{-}49)$$

$$(f_1f_2)(A)=f_1(B)f_2(A) \qquad (3\text{-}50)$$

性质3　若A和B相似，则$f(A)$和$f(B)$也相似。即

$$B=P^{-1}AP\Longrightarrow f(B)=P^{-1}f(A)P \qquad (3\text{-}51)$$

性质4　若A是对角分块矩阵，则$f(A)$亦为对角分块矩阵。即

$$A=A_1\dot{+}A_2\dot{+}\cdots\dot{+}A_s=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

$$\Longrightarrow f(A)=f(A_1)\dotplus f(A_2)\dotplus\cdots\dotplus f(A_s)$$

$$=\begin{pmatrix} f(A_1) & & \\ & f(A_2) & \\ & & \ddots \\ & & & f(A_s) \end{pmatrix} \tag{3-52}$$

性质5 若矩阵A的约当标准形为

$$J=\left(\begin{array}{cccc|cccc} \lambda_1 & 1 & & & & & & \\ & \lambda_1 & 1 & & & & & \\ & & \ddots & \ddots & & & & \\ & & & \ddots\ 1 & & & & \\ & & & \lambda_1 & & & & \\ \hline & & & & \lambda_2 & 1 & & \\ & & & & & \lambda_2 & 1 & \\ & & & & & & \ddots & \ddots \\ & & & & & & & \ddots\ 1 \\ & & & & & & & \lambda_2 \end{array}\right)_{(m+1)\times(m+1)}$$

则 $$f(A)=Pf(J)P^{-1} \tag{3-53}$$

其中 $$f(J)=\begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \tag{3-54}$$

这里

$$D_i=\begin{pmatrix} f(\lambda_i) & \frac{f^1(\lambda_i)}{1!} & \cdots & \frac{f^{(m-1)}(\lambda_i)}{(m-1)!} \\ & \ddots & \ddots & \\ & f(\lambda_i) & \cdots & \frac{f^{(m-2)}(\lambda_i)}{(m-2)!} \\ & & \ddots & \vdots \\ & & & \frac{f'(\lambda_i)}{1!} \\ & & & f(\lambda_i) \end{pmatrix} \quad i=1,\ 2$$

证明 性质1和性质2：若f是多项式，显然(3-48)式，(3-49)式及(3-50)式均成立。

若f为解析函数时，根据定义1，有$f(A)=g(A)$（矩阵多项

式)，则(3-48)式，(3-49)式及(3-50)式亦均成立。

性质3 因为相似矩阵有相同的最小多项式(见第二章习题4(1))，则A和B有相同的最小多项式$\varphi_m(\lambda)$，从而$f(\lambda)$在矩阵A和B上的谱值完全相同，因此可以找到和$f(\lambda)$在A及B上的谱值一致的多项式$g(\lambda)$，则$f(A)=g(A)$，$f(B)=g(B)$。由于$g(A)\sim g(B)$，即$g(B)=P^{-1}g(A)P$（第二章第二节中矩阵多项式的性质2），

故 $f(B)=P^{-1}f(A)P$ 即

$f(A)$和$f(B)$也相似。

性质4 设A的最小多项式为

$$\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s},$$

根据与$f(\lambda)$在A上的谱值相同来确定出多项式$g(\lambda)$，

令 $g(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_{m-1}\lambda^{m-1}$.

其中 $m=\sum\limits_{i=1}^{s}m$

而 $f(A)=g(A)=a_0I+a_1A+a_2A^2+\cdots+a_{m-1}A^{m-1}$

$$=\begin{pmatrix}a_0 & & \\ & a_0 & \\ & & \ddots \\ & & & a_0\end{pmatrix}+\begin{pmatrix}a_1A_1 & & \\ & a_1A_2 & \\ & & \ddots \\ & & & a_1A_s\end{pmatrix}$$

$$+\begin{pmatrix}a_2A_1^2 & & \\ & a_2A_2^2 & \\ & & \ddots \\ & & & a_2A_s^2\end{pmatrix}+\cdots+\begin{pmatrix}a_{m-1}A_1^{m-1} & & \\ & a_{m-1}A_1^{m-1} & \\ & & \ddots \\ & & & a_{m-1}A_s^{m-1}\end{pmatrix}$$

$$=\begin{pmatrix}g(A_1) & & \\ & g(A_2) & \\ & & \ddots \\ & & & g(A_s)\end{pmatrix}=g(A_1)\dot{+}g(A_2)\dot{+}\cdots\dot{+}g(A_s)$$

又因$f(\lambda)$和$g(\lambda)$在A_1，A_2，…，A_s上的谱值相同，

则 $f(A_1)=g(A_1)$，$f(A_2)=g(A_2)$，…，$f(A_s)=g(A_s)$，

故

$$f(A)=\begin{pmatrix} f(A_1) & & & \\ & f(A_2) & & \\ & & \ddots & \\ & & & f(A_s) \end{pmatrix}$$

$$=f(A_1)\dotplus f(A_2)\dotplus\cdots\dotplus f(A_s)$$

性质5：由性质3，性质4及例1很容易证明。

注意：性质5可以推广到A的约当标准形为：

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}_{n\times n}$$ 的情形。

其中

$$J_i=\begin{pmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{i\alpha_i} \end{pmatrix}_{m_i\times m_i},\quad \sum_{i=1}^{s} m_i=n$$

$i=1,\ 2,\cdots,\ s$

$$J_{ij}=\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ & & & \lambda_i \end{pmatrix}_{n_{ij}\times n_{ij}},\quad \sum_{j=1}^{\alpha_i} n_{ij}=m_i$$

$j=1,\ 2,\ \cdots,\ \alpha_i$

则 $f(A)=\mathrm{P}f(J)\mathrm{P}^{-1}$

其中

$$f(J)=\begin{pmatrix} f(J_1) & & & \\ & f(J_2) & & \\ & & \ddots & \\ & & & f(J_s) \end{pmatrix}_{n\times n}$$

$$f(J_i)=\begin{pmatrix} f(J_{i1}) & & & \\ & f(J_{i2}) & & \\ & & \ddots & \\ & & & f(J_{i\alpha_i}) \end{pmatrix}_{m_i\times m_i}$$

$i=1,\ 2\cdots,\ s$

$$f(J_{ij})=\begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \cdots & \frac{f^{(n_{ij}-1)}(\lambda_i)}{(n_{ij}-1)!} \\ & f(\lambda_i) & \cdots & \frac{f^{(n_{ij}-2)}(\lambda_i)}{(n_{ij}-2)!} \\ & & \ddots & \vdots \\ & & & f'(\lambda_i) \\ & & & f(\lambda_i) \end{pmatrix}_{n_{ij}\times n_{ij}}$$

$j=1,2,\cdots,\alpha_i$

由此就得到了求高阶矩阵的矩阵函数的方法：

第一步：求出矩阵A的约当标准形J及变换矩阵P；

第二步：在化$\lambda I-A$为标准形的过程中，求出不变因子，得到A的最小多项式：

$$\varphi_m(\lambda)=d_n(\lambda)=\frac{\psi(\lambda)}{D_{n-1}(\lambda)}=(\lambda-\lambda_1)^{m}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s};$$

第三步：求出$f(\lambda)$在A上的谱值：

$f(\lambda_1)$, $f'(\lambda_1)$, $\cdots$, $f^{(m_1-1)}(\lambda_1)$; $f(\lambda_2)$, $f'(\lambda_2)$, $\cdots$, $f^{(m_2-1)}(\lambda_2)$; $\cdots$, $f(\lambda_s)$, $f'(\lambda_s)$, $\cdots$, $f^{(m_s-1)}(\lambda_s)$;

第四步：写出$f(J)$；

第五步：计算出$Pf(J)P^{-1}$，即得

$f(A)=Pf(J)P^{-1}$

〔**例5**〕 设矩阵

$$A=\begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$$

求矩阵函数A^{20}，e^A及$\sin A$。

解 先把A化成Jordan标准形J，并求出变换矩阵P。因为

$$\lambda I - A = \begin{pmatrix} \lambda-2 & 0 & 0 \\ -1 & \lambda-1 & -1 \\ -1 & 1 & \lambda-3 \end{pmatrix}$$

显然，容易获得特征矩阵$\lambda I - A$的各阶子式的行列式因子。即

$$D_1(\lambda) = 1,$$

$$D_2(\lambda) = \lambda - 2$$

$$D_3(\lambda) = (\lambda-2)^3$$

于是得$\lambda I - A$的不变因子

$$d_1(\lambda) = D_1(\lambda) = 1$$

$$d_2(\lambda) = \frac{D_2(\lambda)}{D_1(\lambda)} = \lambda - 2$$

$$d_3(\lambda) = \frac{D_3(\lambda)}{D_2(\lambda)} = (\lambda-2)^2$$

则$\lambda I - A$的标准形（Smith标准形）为：

$$\begin{pmatrix} 1 & & \\ & \lambda-2 & \\ & & (\lambda-2)^2 \end{pmatrix}$$

所以 $\varphi_m(\lambda) = \dfrac{\psi(\lambda)}{D_2(\lambda)} = (\lambda-2)^2$

由于$\lambda I - A$的初等因子为$\lambda-2$，$(\lambda-2)^2$，得

$$A \sim J = \begin{pmatrix} 2 & & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$$

即 $P^{-1}AP = J$，根据$AP = PJ$求出变换矩阵

$$P=\begin{pmatrix}1&0&1\\1&1&0\\0&1&0\end{pmatrix},\quad P^{-1}=\begin{pmatrix}0&1&-1\\0&0&1\\1&-1&1\end{pmatrix}$$

再计算$f(J)$，由公式（3-54），得

$$f(J)=J^{20}=\begin{pmatrix}2^{20}&&\\&2^{20}&20\times2^{19}\\&&2^{20}\end{pmatrix},\quad e^{J}=\begin{pmatrix}e^2&&\\&e^2&e^2\\&&e^2\end{pmatrix}$$

$$\sin J=\begin{pmatrix}\sin2&&\\&\sin2&\cos2\\&&\sin2\end{pmatrix}$$

又由公式（3-53），即$f(A)=Pf(J)P^{-1}$，得

$$A^{20}=PJ^{20}P^{-1}=\begin{pmatrix}1&0&1\\1&1&0\\0&1&0\end{pmatrix}\begin{pmatrix}2^{20}&0&0\\0&2^{20}&20\times2^{19}\\0&0&2^{20}\end{pmatrix}\begin{pmatrix}0&1&-1\\0&0&1\\1&-1&1\end{pmatrix}$$

$$=\begin{pmatrix}2^{20}&0&0\\20\times2^{19}&2^{20}-20\times2^{19}&20\times2^{19}\\20\times2^{19}&-20\times2^{19}&2^{20}+20\times2^{19}\end{pmatrix}$$

$$e^{A}=Pe^{J}P^{-1}=\begin{pmatrix}1&0&1\\1&1&0\\0&1&0\end{pmatrix}\begin{pmatrix}e^2&0&0\\0&e^2&e^2\\0&0&e^2\end{pmatrix}\begin{pmatrix}0&1&-1\\0&0&1\\1&-1&1\end{pmatrix}$$

$$=\begin{pmatrix}e^2&0&0\\e^2&0&e^2\\e^2&-e^2&2e^2\end{pmatrix}=e^2\begin{pmatrix}1&0&0\\1&0&1\\1&-1&2\end{pmatrix}$$

$$\sin A=P\sin JP^{-1}$$

$$=\begin{pmatrix}1&0&1\\1&1&0\\0&1&0\end{pmatrix}\begin{pmatrix}\sin2&0&0\\0&\sin2&\cos2\\0&0&\sin2\end{pmatrix}\begin{pmatrix}0&1&-1\\0&0&1\\1&-1&1\end{pmatrix}$$

$$=\begin{pmatrix}\sin2&0&0\\\cos2&\sin2-\cos2&\cos2\\\cos2&-\cos2&\sin2+\cos2\end{pmatrix}$$

三、矩阵函数的基本公式

设$f(\lambda)$是任意已知解析函数，矩阵A的最小多项式为：

$$\varphi_m(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$

令 $g(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_k\lambda^k \qquad \left(k=\sum_{i=1}^{s}m_i-1\right)$

使得 $f^{(l)}(\lambda_t)=g^{(l)}(\lambda_t) \quad (t=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ 2,\ \cdots,\ m_t-1)$

即

$$\begin{cases}f(\lambda_1)=a_0+a_1\lambda_1+a_2\lambda_1^2+\cdots+a_k\lambda_1^k\\ f'(\lambda_1)=a_1+2a_2\lambda_1+\cdots+ka_k\lambda_1^{k-1}\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ f^{(m_1-1)}(\lambda_1)=a_{m_1-1}(m_1-1)!+a_{m_1}m_1!\lambda_1+\cdots\\ \qquad +a_k\dfrac{k!}{(k-m_1+1)!}\lambda_1^{k-(m_1-1)}\end{cases}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\begin{cases}f(\lambda_s)=a_0+a_1\lambda_s+a_2\lambda_s^2+\cdots+a_k\lambda_s^k\\ f'(\lambda_s)=a_1+2a_2\lambda_s+\cdots+ka_k\lambda_s^{k-1}\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ f^{(m_s-1)}(\lambda_s)=a_{m_s-1}(m_s-1)!+a_{m_s}m_s!\lambda_s+\cdots\\ \qquad +a_k\dfrac{k!}{(k-m_s+1)!}\lambda_s^{k-(m_s-1)}\end{cases}$$

解上方程组，求出$a_0,a_1,\cdots,a_k$。显然，a_i是$f(\lambda_1),f'(\lambda_1),\cdots,$

$f^{(m_1-1)}(\lambda_1)$；…；$f(\lambda_s)$，$f'(\lambda_s)$，…，$f^{(m_s-1)}(\lambda_s)$的线性函数，其系数是λ_i的函数。把$a_i(i=0, 1, 2, \cdots, k)$代回$g(\lambda)$，得到

$$g(\lambda)=Z_{11}f(\lambda_1)+Z_{12}f'(\lambda_1)+\cdots+Z_{1m_1}f^{(m_1-1)}(\lambda_1)$$
$$+\cdots+Z_{s1}f(\lambda_s)+Z_{s2}f'(\lambda_s)+\cdots+Z_{sm_s}f^{(m_s-1)}(\lambda_s)$$

其中

Z_{11}，Z_{12}，…，Z_{1m_1}；…；Z_{s1}，Z_{s2}，…，Z_{sm_s}都是λ，λ_i的函数。

故得**矩阵函数的基本公式**：

$$f(A)=g(A)=Z_{11}(A)f(\lambda_1)+Z_{12}(A)f'(\lambda_1)+\cdots$$
$$+Z_{1m_1}(A)f^{(m_1-1)}(\lambda_1)+\cdots+Z_{s1}(A)f(\lambda_s)+Z_{s2}(A)f'(\lambda_s)+\cdots$$
$$+Z_{sm_s}(A)f^{(m_s-1)}(\lambda_s) \qquad (3\text{-}55)$$

其中Z_{ij}只与矩阵A有关，与函数$f(\lambda)$无关。由于A的谱和$f(\lambda)$在A上的谱值$f(\lambda_1)$，$f'(\lambda_1)$，…，$f^{(m_1-1)}(\lambda_1)$；…；$f(\lambda_s)$，$f'(\lambda_s)$，…，$f^{(m_s-1)}(\lambda_s)$是唯一的，所以$f(A)$表示成（3-55）式的形式也是唯一的。又因为Z_{ij}与$f(\lambda)$无关，即是不论$f(\lambda)$为任何解析函数，Z_{ij}总是唯一确定的。故可取一些特殊的函数$f_1(\lambda)$，$f_2(\lambda)$，…，以便求出Z_{ij}来。

公式（3-55）在系统理论中应用是非常方便的。

〔**例6**〕 计算A^{100}，

其中 $$A=\begin{pmatrix} 5 & -4 \\ 4 & -3 \end{pmatrix}$$

解 设$f(\lambda)=\lambda^{100}$，因为A的最小多项式为：

$$\varphi_m(\lambda)=(\lambda-1)^2$$

由(3-55)式，得 $f(A)=Z_1(A)f(1)+Z_2(A)f'(1)$ （3-56）

由于Z_1，Z_2与f无关，故可取

$$f_1(\lambda)=1, \qquad f'(\lambda)=0$$

$$f_2(\lambda)=\lambda-1,\quad f_2(1)=0,\quad f_2'(\lambda)=1$$

代入（3-56）式，得

$$\begin{cases} I=Z_1(A) \\ A-I=Z_2(A) \end{cases}$$

于是

$$f(A)=f(1)I+f'(1)(A-I)$$ 因为

$f(\lambda)=\lambda^{100}$，得

$$f(1)=1,\quad f'(1)=100$$

故$f(A)=A^{100}=I+100(A-I)=100A-99I$

$$=\begin{pmatrix} 401 & -400 \\ 400 & -399 \end{pmatrix}$$

四、矩阵函数的幂级数表示

回忆本章第一节中所提出的矩阵级数问题，特别是矩阵幂级数对研究矩阵理论和在实用上都是很重要的。前面对矩阵级数的收敛和发散的概念已给了定义，这里我们还需要继续阐明一些基本概念及性质。

定义2　若矩阵级数$\sum_{m=1}^{\infty} A_m$中的n^2个数项级数都是绝对收敛的，则称矩阵级数$\sum_{m=1}^{\infty} A_m$是绝对收敛的。

由此定义及数学分析中的结果，相应可得下列性质：

性质1　若矩阵级数$\sum_{m=1}^{\infty} A_m$绝对收敛，则它一定收敛，且任意调换各项次序后所得的新级数仍是收敛的，而且和不变。

性质2　矩阵级数$\sum_{m=1}^{\infty} A_m$绝对收敛的充要条件是

$\sum_{m=1}^{\infty}\|A_m\|$收敛。

证明 根据范数的等价性，必存在正常数$c_1<c_2$满足

$$c_1\|A_m\|\leqslant\|A_m\|_1\leqslant c_2\|A_m\| \tag{3-57}$$

必要性： 设$\sum_{m=1}^{\infty}A_m$绝对收敛，则$A_m=\left[a_{ij}^{(m)}\right]_{n\times n}$的元素所构成的$n^2$个数项级数$\sum_{m=1}^{\infty}|a_{ij}^{(m)}|$ ($i, j=1, 2, \cdots, n$) 收敛。于是存在正数M，使得

$$\sum_{m=1}^{N}|a_{ij}^{(m)}|<M \quad (N\geqslant 1;\ i,\ j=1,\ 2,\cdots,n)$$

由于 $\|A_m\|_1=\max\limits_{j}\sum_{i=1}^{n}|a_{ij}^{(m)}|$，因此有

$$\sum_{m=1}^{N}\|A_m\|_1\leqslant\sum_{m=1}^{N}\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}^{(m)}|\leqslant n^2M$$

而 $\left\{\sum_{m=1}^{N}\|A_m\|_1\right\}$单调增，

故 $\sum_{m=1}^{\infty}\|A_m\|_1$收敛，由(3-57)式即知

$$\sum_{m=1}^{\infty}\|A_m\|\text{也收敛。}$$

充分性： 若$\sum_{m=1}^{\infty}\|A_m\|$收敛，据(3-57)式级数

$\sum_{m=1}^{\infty}\|A_m\|_1$也收敛，由于

$$\|A_m\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}^{(m)}|$$

故有 $|a_{ij}^{(m)}| \leqslant \|A_m\|_1$ $\quad (i, j=1, 2, \cdots, n)$

于是由比较判别法，得知 $\sum_{m=1}^{\infty} A_m$ 中的 n^2 个数项级数

$$\sum_{m=1}^{\infty} |a_{ij}^{(m)}| \qquad (i, j=1, 2, \cdots, n)$$

都收敛。

再由定义 2，即知级数 $\sum_{m=1}^{\infty} A_m$ 绝对收敛。证毕。

有了上述这识，就可以考虑将矩阵函数 $f(A)$ 定义为绝对收敛的矩阵幂级数的和，即定义

$$f(A) = \sum_{k=0}^{\infty} c_k A^k$$

定理1 设矩阵 A 的最小多项式为：

$$\varphi_m(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$

$f(\lambda)$ 及 $f_n(\lambda)$ $\quad (n=1, 2, \cdots)$ 都是在包含 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 的开域上的解析函数。若

$$\lim_{s\to\infty} f_s^{(l)}(\lambda_k) = f^{(l)}(\lambda_k) \tag{3-58}$$

$$(k=1, 2\cdots, s;\ l=0, 1, 2, \cdots, m_k-1)$$

则
$$\lim_{s\to\infty} f_n(A) = f(A) \tag{3-59}$$

证明 设 (3-58) 式成立，由(3-55)式，知

$$f_n(A) = Z_{11}(A)f_n(\lambda_1) + Z_{12}(A)f'_n(\lambda_1) + \cdots$$
$$+ Z_{1m_1}(A)f_n^{(m_1-1)}(\lambda_1) + \cdots + Z_{s1}(A)f_n(\lambda_s) + Z_{s2}(A)f'_n(\lambda_s)$$
$$+ \cdots + Z_{sm_s}(A)f_n^{(m_s-1)}(\lambda_s)$$

则 $\lim\limits_{n\to\infty} f_n(A) = Z_{11}(A) \lim\limits_{n\to\infty} f_n(\lambda_1) + Z_{12}(A) \lim\limits_{n\to\infty} f'_n(\lambda_1)$

$$+\cdots+Z_{1m_1}(A)\lim_{n\to\infty}f_n^{(m_1-1)}(\lambda_1)+\cdots+Z_{s1}(A)\lim_{n\to\infty}f_n(\lambda_s)$$

$$+Z_{s_2}(A)\lim_{n\to\infty}f_n'(\lambda_s)+\cdots+Z_{sm_s}\lim_{n\to\infty}f_n^{(m_s-1)}(\lambda_s)$$

$$=Z_{11}(A)f(\lambda_1)+_{12}(A)f'(\lambda_1)+\cdots+Z_{1m_1}(A)f^{(m_1-1)}(\lambda_1)+\cdots$$

$$+Z_{s1}(A)f(\lambda_s)+Z_{s2}(A)f'(\lambda_s)+\cdots+Z_{sms}(A)f^{(m_s-1)}(\lambda_s)$$

故 $\lim_{n\to\infty}f_n(A)=f(A)$

定理2 设幂级数$\sum_{k=0}^{\infty}c_k\lambda^k$的收敛半径为$r>0$，当矩阵$A$的全部特征值都在半径为$r$的圆内时，则矩阵幂级数

$$\sum_{k=0}^{\infty}c_kA^k$$

绝对收敛，，且当$f(\lambda)=\sum_{k=0}^{\infty}c_k\lambda^k$时，有

$$f(A)=\sum_{k=0}^{\infty}c_kA^k$$

证明 设 $f_n(\lambda)=\sum_{k=0}^{n}c_k\lambda^k$，且

$$\lim_{n\to\infty}f_n(\lambda)=f(\lambda)$$

由幂级数性质，知

$f_n^{(l)}(\lambda)$对于$|\lambda|<r$绝对收敛。

且有

$\lim_{n\to\infty}f_n^{(l)}(\lambda_k)=f^{(l)}(\lambda_k)$，对于$|\lambda_k|<r$

$(k=1, 2, \cdots, s;\ l=0, 1, 2, \cdots, m_k-1)$

由定理 1，得

$$\lim_{s\to\infty}f_n(A)=\sum_{k=0}^{\infty}c_kA^k=f(A)$$ 证毕。

〔例7〕 对于

$$e^{\lambda}=1+\lambda+\frac{1}{2!}\lambda^2+\cdots+\frac{1}{k!}\lambda^k+\cdots,\quad |\lambda|<\infty$$ 在 $|\lambda|<\infty$ 内，上面级数可以求任意阶导数，它们所构成的级数也收敛。则对任意矩阵A，当其特征值$\lambda_t(t=1,2,\cdots,s)$，都在复平面上时，均有

$$e^A=I+A+\frac{A^2}{2!}+\cdots+\frac{A^k}{k!}+\cdots \qquad (|\lambda_t|<\infty;\ t=1,\ 2,\ \cdots,\ s)$$

同理对于

$$(1-\lambda)^{-1}=1+\lambda+\lambda^2+\cdots+\lambda^k+\cdots,\quad |\lambda|<1$$

有 $$(I-A)^{-1}=I+A+A^2+\cdots+A^k+\cdots$$

$$(|\lambda_t|<1;\ t=1,\ 2,\ \cdots,\ s)$$

对于 $$\sin\lambda=\lambda-\frac{\lambda^3}{3!}+\frac{\lambda^5}{5!}-\cdots$$

$$+(-1)^{k-1}\frac{\lambda^{2k-1}}{(2k-1)!}+\cdots \quad (|\lambda_t|<\infty)$$

有$$\sin A=A-\frac{A^3}{3!}+\frac{A^5}{5!}-\cdots+(-1)^{k-1}\frac{A^{2k-1}}{(2k-1)!}+\cdots$$

$$(|\lambda_t|<\infty;\ t=1,2,\cdots s)$$

对于 $$\ln(1+\lambda)=\lambda-\frac{\lambda^2}{2}+\frac{\lambda^3}{3}-\frac{\lambda^4}{4}+\cdots+(-1)^{k-1}\frac{1}{k}\lambda^k+\cdots(|\lambda|<1)$$

有 $$\ln(I+A)=A-\frac{A^2}{2}+\frac{A^3}{3}-\frac{A^4}{4}+\cdots+(-1)^{k-1}\frac{1}{k}A^k+\cdots+(|\lambda_t|<1,\ t=1,2,\cdots s)$$

习　题　三

1. 设函数阵矩

$$A(t)=\begin{pmatrix} \sin t & \cos t & t \\ \dfrac{\sin t}{t} & e^t & t^2 \\ 1 & 0 & t^3 \end{pmatrix} \qquad (t\neq 0)$$

试计算：$\lim\limits_{t\to 0}A(t)$，$\dfrac{d}{dt}A(t)$，$\dfrac{d^2}{dt^2}A(t)$，$\dfrac{d}{dt}\det A(t)$，$\det\dfrac{d}{dt}A(t)$，

$\dfrac{d}{dt}A^{-1}(t)$

2. 设函数矩阵

$$A(x)=\begin{pmatrix} e^{2x} & xe^x & 1 \\ e^{-x} & 2e^{2x} & 0 \\ 3x & 0 & 0 \end{pmatrix}$$

试计算：$\int_0^1 A(x)dx$ 和 $\int A(x)dx$

3. 对任意n阶矩阵 $A=[a_{ij}]_{n\times n}$，定义

$\|A\|=\max\limits_{1\leq i,\,j\leq n}|a_{ij}|$为$A$的范数（定义2中的条件(4)应改为

$\|AB\|\leqslant n\|A\|\,\|B\|$，$B\in F^{n\times n}$）证明：

(1) 矩阵序列$\{A_m\}\to 0$ $(m\to\infty)$的充要条件是数列$\{\|A_m\|\}\to 0$ $(m\to\infty)$，即

$$\lim_{m\to\infty}A_m=0\Longleftrightarrow\lim_{m\to\infty}\|A_m\|=0$$

(2) 对k个n阶矩阵A_1，A_2，…，A_k，恒有

$$\|A_1A_2\cdots A_k\|\leqslant n^{k-1}\|A_1\|\cdot\|A_2\|\cdots\|A_k\|$$

4. (1) 设矩阵

$$A=\begin{pmatrix}1 & -4 & -1 & -4\\ 2 & 0 & 5 & -4\\ -1 & 1 & -2 & 3\\ -1 & 4 & -1 & 6\end{pmatrix}$$

求$\|A\|_1$和$\|A\|_\infty$

(2) 设矩阵

$$A=\begin{pmatrix}\frac{1}{2} & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}$$

求$\|A\|_2$，$\|A\|_\infty$，$\|A^{-1}\|_\infty$

(3) 在R^n（或C^n）上，证明：

$$\|X\|_1\geqslant\|X\|_2\geqslant\|X\|_\infty\geqslant\frac{1}{n}\|X\|_1$$

(4) 证明：

$\|A\|=n\max\limits_{i,j}|a_{ij}|$是矩阵$A=[a_{ij}]_{n\times n}$的范数。

5．设矩阵 $A=\begin{pmatrix}1 & 4\\ 3 & 2\end{pmatrix}$， $B=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}$

求e^A，e^B，$\sin B$，$\cos B$

6．设矩阵

$$A=\begin{pmatrix}0 & -1\\ 4 & 4\end{pmatrix}$$

求 $\sin^{-1}(\frac{1}{4})A$

7．证明：对任意矩阵A，下列关系式恒成立。

$\sin^2A+\cos^2A=I$

$e^{A+2\pi iI}=e^A$

$\sin(A+2\pi I)=\sin A$

8. 设矩阵

$$A=\begin{pmatrix}0 & 1\\ -1 & 2\end{pmatrix},$$

求$A^{100}+3A^{23}+A^{20}$

9. 设矩阵

$$A=\begin{pmatrix}2 & 2 & 1\\ 1 & 3 & 1\\ 1 & 2 & 2\end{pmatrix}$$

求A^{1000}

10. 设矩阵

$$A=\begin{pmatrix}0 & 0 & -2\\ 0 & 1 & 0\\ 1 & 0 & 3\end{pmatrix},\qquad B=\begin{pmatrix}3 & 0 & 0 & 0\\ 0 & -2 & 1 & 0\\ 0 & 0 & -2 & 1\\ 0 & 0 & 0 & -2\end{pmatrix}$$

$$C=\begin{pmatrix}0 & 1\\ -2 & -3\end{pmatrix},$$

求矩阵函数e^{At}，e^{Bt}和e^{Ct}

11. 设A，B都是n阶可换矩阵，即

$AB=BA$

证明：

(1) $e^{(A+B)t}=e^{At}e^{Bt}$；

(2) $e^{A+iB}=e^{A}(\cos B+i\sin B)$

第四章　微分方程的矩阵分析解法

对于线性微分方程可以用各种各样的方法求解，每一种方法，都有其各自的特点，适应不同的计算需要，在科学技术的应用中，求解线性微分方程，经常引用Jordan标准形、矩阵函数以及计算机求数值解等方法。

本章对线性微分方程的解的结构和求解方法进行了必要的分析，也对科学技术研究中常遇见的变系数线性矩阵微分方程及矩阵Riccati方程的求解问题，作了必要的论述。

第一节　线性微分方程系统的解的结构

在实际问题，例如状态空间方程的讨论中，需要对以下形式的矩阵微分方程求解：

$$\dot{X}=A(t)X+B(t)u \tag{4-1}$$

其中X是n维向量，u是r维向量；

$$A(t)=[a_{ij}(t)]_{n\times n};$$

$$B(t)=[b_{ij}(t)]_{n\times r}$$

$A(t)$，$B(t)$在区间$[t_0, t_1]$上，函数是绝对可积的。

可以证明，对给定的t_0和$X(t_0)$，方程(4-1)在区间$[t_0, t_1]$上有唯一解。

一、基础解系和Wronskian行列式

基础解系和Wronskian行列式是常微分方程理论中两个重要的基本概念。

考虑(4-1)式的齐次形式

$$\dot{X}=A(t)X \tag{4-2}$$

如果X_1，X_2，…，X_n是方程（4-2）的一组线性无关解，则它的其它任一解可以表为它的线性组合。只要初始条件一确定，这个解就可以唯一确定。

一般地，方程(4-2)的全部解将构成一个n维向量的解空间，其中任一组n个线性无关的解，就称为它的一个基础解系。

判断n个向量X_1，X_2，…，X_n是否为方程(4-2)的线性无关解（即基础解系）的充要条件（将由后面定理4得出）是

$$W=\begin{vmatrix} x_{11} & x_{12}\cdots x_{1n} \\ x_{21} & x_{22}\cdots x_{2n} \\ \cdots & \cdots\cdots \\ x_{n1} & x_{n2}\cdots x_{nn}\end{vmatrix}\neq 0$$

这个行列式称为Wronskian（朗斯基）行列式。

定理1 若n个可微函数$X_1(t)$，$X_2(t)$，…，$X_n(t)$在区间$[t_0, t_1]$上每一点都是线性相关的，则朗斯基行列式$W(X_1, X_2, \cdots, X_n)$在区间$[t_0, t_1]$上恒为零。

证明 因为X_1，X_2，…，X_n在区间$[t_0, t_1]$上线性相关，则有不全为零的常数c_i($i=1, 2, \cdots, n$)，使得

$$C_1X_1(t)+C_2X_2(t)+\cdots+C_nX_n(t)=0 \tag{4-3}$$

对（4-3）式逐次微分可得

$$C_1\dot{X}_1(t)+C_2\dot{X}_2(t)+\cdots+C_n\dot{X}_n(t)=0$$
$$C_1\ddot{X}_1(t)+C_2\ddot{X}_2(t)+\cdots+C_n\ddot{X}_n(t)=0$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$C_1\overset{(n-1)}{X_1}(t)+C_2\overset{(n-1)}{X_2}(t)+\cdots+C_n\overset{(n-1)}{X_n}(t)=0$$

取$t=t_2$，且$t_0 \leqslant t_2 \leqslant t_1$，代入(4-3)式及它的微分式，则得

$$\begin{cases} C_1X_1(t_2)+C_2X_2(t_2)+\cdots+C_nX_n(t_2)=0 \\ C_1\dot{X}_1(t_2)+C_2\dot{X}_2(t_2)+\cdots+C_n\dot{X}_n(t_2)=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ C_1\overset{(n-1)}{X_1}(t_2)+C_2\overset{(n-1)}{X_2}(t_2)+\cdots+C_n\overset{(n-1)}{X_n}(t_2)=0 \end{cases}$$

其中$\overset{(j-1)}{X_i}(t_2)$，$(i,\ j=1,\ 2,\ \cdots,\ n)$为常数，因为$c_i$不全为零，所以只有当其系数行列式，即

$$\begin{vmatrix} X_1(t_2) & X_2(t_2)\cdots & X_n(t_2) \\ \dot{X}_1(t_2) & \dot{X}_2(t_2)\cdots & \dot{X}_n(t_2) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \overset{(n-1)}{X_1}(t_2) & \overset{(n-1)}{X_2}(t_2)\cdots & \overset{(n-1)}{X_n}(t_2) \end{vmatrix}=0$$

时，上面方程组对c_i才有非零解。又因为t_2在区间$[t_0,\ t_1]$内是任意的，于是，对于任何$t\in[t_0,\ t_1]$都有函数组$X_1(t)$，$X_2(t)$，$\cdots$，$X_n(t)$的朗斯基行列式

$$W(X_1,X_2,\cdots,X_n)=\begin{vmatrix} X_1(t) & X_2(t) & \cdots X_n(t) \\ \dot{X}_1(t) & \dot{X}_2(t) & \cdots \dot{X}_n(t) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \overset{(n-1)}{X_1}(t) & \overset{(n-1)}{X_2}(t)\cdots & \overset{(n-1)}{X_n}(t) \end{vmatrix}=0$$

至此证毕。但应注意到：逆定理并不成立。

例如：考虑

$$X_1(t)=t^3,\quad X_2(t)=|t|^3$$

则有

$$W(X_1,\ X_2)=\begin{vmatrix} t^3 & \pm t^3 \\ 3t^2 & \pm 3t^2 \end{vmatrix}=0$$

显然，$X_1(t)$，$X_2(t)$在$[-1,\ 1]$上是线性无关的。

定理2 若n个可微函数$X_1(t)$，$X_2(t)$，…，$X_n(t)$的朗斯基行列式$W(X_1, X_2, \cdots, X_n)$在区间$[t_0, t_1]$上为零，而$n-1$个函数$X_1(t)$，$X_2(t)$，…，$X_{n-1}(t)$的朗斯基行列式在整个区间$[t_0, t_1]$上不为零，则n个函数$X_1(t)$，$X_2(t)$，…，$X_n(t)$在区间$[t_0, t_1]$上是线性相关的。

证明 由假设，在区间$[t_0, t_1]$上有$W(X_1, X_2, \cdots, X_{n-1})\neq 0$，则对任一$t\in[t_0,t_1]$都存在一组$n-1$个函数$C_1(t)$，$C_2(t)$，…，$C_{n-1}(t)$，使得

$$\begin{cases} C_1X_1(t)+C_2X_2(t)+\cdots+C_{n-1}X_{n-1}(t)=X_n(t) \\ C_1\dot{X}_1(t)+C_2\dot{X}_2(t)+\cdots+C_{n-1}\dot{X}_{n-1}(t)=\dot{X}_n(t) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ C_1\overset{(n-2)}{X_1}(t)+\dot{C}_2\overset{(n-2)}{X_2}(t)+\cdots+C_{n-1}\overset{(n-2)}{X}_{n-1}(t)=\overset{(n-2)}{X_n}(t) \end{cases} \tag{4-4}$$

其中C_i $(i=1, 2, \cdots, n-1)$ 假定可微，且取

$$C_1\overset{(n-1)}{X_1}(t)+C_2\overset{(n-1)}{X_2}(t)+\cdots+C_{n-1}\overset{(n-1)}{X}_{n-1}(t)=\overset{(n-1)}{X_n}(t)+C \tag{4-5}$$

在(4-4)式及(4-5)式的两边分别乘以$W(X_1,X_2,\cdots,X_n)$中X_n，$\dot{X}_n$，…，$\overset{(n-2)}{X_n}$，$\overset{(n-1)}{X_n}$的代数余子式后，将所得诸方程相加，并注意到左端每个行列式中有两列相同，则得

$$0=W(X_1, X_2, \cdots, X_n)+CW(X_1, X_2, \cdots, X_{n-1})$$

由于

$$W(X_1, X_2, \cdots, X_n)=0, \quad W(X_1, X_2, \cdots, X_{n-1})\neq 0$$

于是有$c=0$

因此，方程(4-5)可写为

$$C_1\overset{(n-1)}{X_1}(t)+C_2\overset{(n-1)}{X_2}(t)+\cdots+C_{n-1}\overset{(n-1)}{X}_{n-1}(t)=\overset{(n-1)}{X_n}(t)$$

若证得C_1，C_2，…，C_{n-1}为常数，即可得出结论：X_1，X_2，…，X_n在区间$[t_0, t_1]$上是线性相关的。为此，首

先将方程组(4-4)中第一个方程对t微分，减去第二个方程即得到

$$\dot{C}_1X_1(t)+\dot{C}_2X_2(t)+\cdots+\dot{C}_{n-1}X_{n-1}(t)=0$$

再微分方程组(4-4)第二个方程，减去第三个方程，同理可得

$$\dot{C}_1\dot{X}_1(t)+\dot{C}_2\dot{X}_2(t)+\cdots+\dot{C}_{n-1}\dot{X}_{n-1}(t)=0$$

类似地，可得到：

$$\dot{C}_1\ddot{X}_1(t)+\dot{C}_2\ddot{X}_2(t)+\cdots+\dot{C}_{n-1}\ddot{X}_{n-1}(t)=0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\dot{C}_1\overset{(n-2)}{X_1}(t)+\dot{C}_2\overset{(n-2)}{X_2}(t)+\cdots+\dot{C}_{n-1}\overset{(n-2)}{X_{n-1}}(t)=0$$

于是有

$$\begin{pmatrix} X_1(t) & X_2(t)\cdots X_{n-1}(t) \\ \dot{X}_1(t) & \dot{X}_2(t)\cdots\dot{X}_{n-1}(t) \\ \cdots\cdots & \cdots\cdots\cdots\cdots \\ \overset{(n-2)}{X_1}(t) & \overset{(n-2)}{X_2}(t)\cdots\overset{(n-2)}{X_{n-1}}(t) \end{pmatrix}\begin{pmatrix} \dot{C}_1 \\ \dot{C}_2 \\ \vdots \\ \dot{C}_{n-1} \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

根据假设：

$$W(X_1,\ X_2,\ \cdots,\ X_{n-1})\neq 0$$

所以

$$\dot{C}_1=\dot{C}_2=\cdots=\dot{C}_{n-1}\equiv 0$$

于是

C_i为常数（$i=1,\ 2,\ \cdots,\ n-1$），从而有

$$C_1X_1(t)+C_2X_2(t)+\cdots+C_{n-1}X_{n-1}(t)=X_n(t)$$

这就证得

X_1，X_2，…，X_n在区间$[t_0,\ t_1]$上线性相关。

二、关于线性微分方程解的几个定理

定理3 设$X_1(t)$，$X_2(t)$，…，$X_n(t)$是线性微分方程

$$L(t)=\overset{(n)}{X}+a_1(t)\overset{(n-1)}{X}+\cdots+a_{n-1}(t)\dot{X}+a_n(t)X=0$$

的 n 个解，其中$a_i(t)$在区间$[t_0, t_1]$中绝对可积。若 Wronskian行列式$W(X_1, X_2, \cdots, X_n)$在$t=t_2$ $(t_0\leqslant t_2\leqslant t_1)$时为零，则$X_1(t)$，$X_2(t)$，…，$X_n(t)$，$t\in[t_0, t_1]$线性相关，而且$W(X_1, X_2, \cdots, X_n)=0$，在整个区间$[t_0, t_1]$上成立。

证明 根据假设，存在不全为零的C_1，C_2，…，C_n，使得

$$\begin{cases} C_1X_1(t_2)+C_2X_2(t_2)+\cdots+C_nX_n(t_2)=0 \\ C_1\dot{X}_1(t_2)+C_2\dot{X}_2(t_2)+\cdots+C_n\dot{X}_n(t_2)=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ C_1\overset{(n-1)}{X_1}(t_2)+C_2\overset{(n-1)}{X_2}(t_2)+\cdots+C_n\overset{(n-1)}{X_n}(t_2)=0 \end{cases}$$

这一组方程，可以看作是$L(t)=0$的解：

$$Y(t)=C_1X_1(t)+C_2X_2(t)+\cdots+C_nX_n(t)$$

满足

$$Y(t_2)=\dot{Y}(t_2)=\ddot{Y}(t_2)=\cdots=\overset{(n-1)}{Y}(t_2)=0$$

从另一面讲，恒等于零的函数也是$L(t)=0$的解，而且此函数的导数全为零。由线性微分方程解的唯一性定理，可得到$Y(t)=0$。所以

$$C_1X_1(t)+C_2X_2(t)+\cdots+C_nX_n(t)=0$$

对任一$t\in[t_0, t_1]$成立，又因$C_i(i=1, 2, \cdots, n)$不全为零，所以$X_1(t)$，$X_2(t)$，…，$X_n(t)$线性相关。再由定理 1，可得到$W(X_1, X_2, \cdots, X_n)=0$，对$t\in[t_0, t_1]$均成立。证毕。

根据定理 3 及定理 1 又可得到定理 4 。

定理4 若线性微分方程

$$L(t)=\overset{(n)}{X}+a_1(t)\overset{(n-1)}{X}+\cdots+a_{n-1}(t)\dot{X}+a_n(t)X=0$$

的解$X_1(t)$，$X_2(t)$，…，$X_n(t)$在区间〔t_0，t_1〕上是线性无关的，其中$a_i(t)$ ($i=1$，2，…，n) 在区间〔t_0，t_1〕上是绝对可积的，则在整个区间〔t_0，t_1〕上Wronskian行列式$W(X_1, X_2, \cdots, X_n) \neq 0$

证明是显然的。

现在把定理4中的微分方程用矩阵微分方程 $\dot{X}=A(t)X$ 的形式写出：即令

$X_1=X$，$X_2=\dot{X}$，…，$X_n=\overset{(n-1)}{X}$，便得到

$$\begin{pmatrix}\dot{X}_1\\ \dot{X}_2\\ \vdots\\ \dot{X}_n\end{pmatrix}=\begin{pmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ -a_n(t) & -a_{n-1}(t) & -a_{n-2}(t) & \cdots & -a_1(t)\end{pmatrix}\begin{pmatrix}X_1\\ X_2\\ \vdots\\ X_n\end{pmatrix}$$

即 $\dot{X}=A(t)X$

显然，若$X_i(t)$是$L(t)=0$的一个解，则

$$X_i=[X_i,\ \dot{X}_i,\ \ddot{X}_i,\ \cdots,\ \overset{(n-1)}{X_i}]^{T}$$

便是$\dot{X}=A(t)X$的一个解，反之也成立。

方程 $\dot{X}=A(t)X$

的全部解形成一个n维向量空间。由于有

$$X_i=\begin{pmatrix}X_{1i}\\ X_{2i}\\ \vdots\\ X_{ni}\end{pmatrix}=\begin{pmatrix}X_i\\ \dot{X}_i\\ \vdots\\ \overset{(n-1)}{X_i}\end{pmatrix}$$

则 Wronskian行列式

$$W(X_1, X_2, \cdots, X_n)=\begin{vmatrix}X_1 & X_2 & \cdots & X_n\\ \dot{X}_1 & \dot{X}_2 & \cdots & \dot{X}_n\\ \cdots & \cdots & \cdots & \cdots\\ \overset{(n-1)}{X_1} & \overset{(n-1)}{X_2} & \cdots & \overset{(n-1)}{X_n}\end{vmatrix}$$

$$
=\begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix}
$$

其后一表达式，就是我们开始时对矩阵微分方程

$$
\dot{X}=A(t)X
$$

所定义的Wronskian行列式。所以定理4只不过是说明X_1，X_2，…，X_n是线性无关，其中

$$
X_i=[x_{1i},\ x_{2i},\ \cdots,\ x_{ni}]^{\mathrm{T}}=[X_i,\ \dot{X}_i,\ \cdots,\ \overset{(n-1)}{X_i}]^{\mathrm{T}}
$$

从而是解空间的一组基底（基础解系）的充要条件是：

$$
W(t_0)=\begin{vmatrix} X_1(t_0) & X_2(t_0) & \cdots & X_n(t_0) \\ \dot{X}_1(t_0) & \dot{X}_2(t_0) & \cdots & \dot{X}_n(t_0) \\ \cdots & \cdots & \cdots & \cdots \\ \overset{(n-1)}{X_1}(t_0) & \overset{(n-1)}{X_2}(t_0) & \cdots & \overset{(n-1)}{X_n}(t_0) \end{vmatrix}
$$

$$
=\begin{vmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{vmatrix} \neq 0
$$

定理5 若线性微分方程

$$
L(t)=\overset{(n)}{X}+a_1(t)\overset{(n-1)}{X}+\cdots+a_n(t)X=0
$$

有一组n个线性无关的解$X_1(t)$，$X_2(t)$，…，$X_n(t)$，其中系数$a_i(t)$ ($i=1$，2，…，n) 在区间$[t_0,\ t_1]$上绝对可积，则$L(t)=0$的一般解为

$$
X(t)=c_1X_1(t)+c_2X_2(t)+\cdots+c_nX_n(t)
$$

其中c_i($i=1$，2，…，n)为任意常数。

证明 由于$X_1(t)$，$X_2(t)$，…，$X_n(t)$线性无关，则

Wronskian行列式$W(X_1, X_2, \cdots, X_n) \neq 0$。若令$Z(t)$是$L(t)=0$的任一解，则由

$$\begin{cases} c_1X_1(t_1)+c_2X_2(t_1)+\cdots+c_nX_n(t_1)=Z(t_1) \\ c_1\dot{X}_1(t_1)+c_2\dot{X}_2(t_1)+\cdots+c_n\dot{X}_n(t_1)=\dot{Z}(t_1) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ c_1\overset{(n-1)}{X_1}(t_1)+c_2\overset{(n-1)}{X_2}(t_1)+\cdots+c_n\overset{(n-1)}{X_n}(t_1)=\overset{(n-1)}{Z}(t_1) \end{cases}$$

可解得不全为零的常数$C_i(i=1, 2, \cdots, n)$

据此，设

$$Y(t)=C_1X_1(t)+C_2X_2(t)+\cdots+C_nX_n(t)$$

则有 $L[Y(t_1)]=0$

且 $Y(t_1)=Z(t_1)$，$\dot{Y}(t_1)=\dot{Z}(t_1)$，$\cdots$，$\overset{(n-1)}{Y}(t_1)=\overset{(n-1)}{Z}(t_1)$

因此，$Y(t)$和$Z(t)$是给定微分方程满足同一初始条件的解。根据解的唯一性，得

$$Y(t)\equiv Z(t)$$

这就证明了微分方程的任一个解$Z(t)$可表为：

$$Z(t)=\sum_{i=1}^{n}c_iX_i(t)$$

由于它含有n个任意常数c_i，所以是一般解。

需要说明的是：

非齐次矩阵微分方程 $\dot{X}(t)=A(t)X(t)+B(t)U(t)$

$$X(t_0)=X_0 \tag{4-6}$$

的解，是由齐次矩阵微分方程的解与方程(4-6)的一个特解组成。

第二节　常系数线性微分方程系统

（线性非时变系统）

一、常系数线性齐次方程

以讨论自由动力系统的运动为例，可写出矩阵微分方程

$$\frac{\mathrm{d}}{\mathrm{d}t}X(t)=AX(t) \tag{4-7}$$

$$X(0)=X_0$$

其中

$$X(t)=\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix};\quad A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix};$$

$$X_0=\begin{pmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{pmatrix}$$

鉴于　$\mathrm{e}^A=I+A+\frac{1}{2!}A+\cdots+\frac{1}{n!}A^n+\cdots$

可以证明　$X(t)=\mathrm{e}^{At}K$

为方程(4-7)的解。其中K为任意常向量

$K=(K_1,\ K_2,\ \cdots,\ K_n)^{\mathrm{T}}$

这个解类似于一阶线性齐次方程，显然，只须将

$$X(t)=\mathrm{e}^{At}K=\left(I+At+\frac{t^2}{2!}A^2+\cdots+\frac{t^n}{n!}A^n+\cdots\right)K$$

代入(4-7)式右侧，即可得到

$$AX(t)=(A+A^2t+\frac{t^2}{2!}A^3+\cdots+\frac{t^n}{n!}A^{n+1}+\cdots)K \tag{4-8}$$

而左侧同时也有：

$$\frac{d}{dt}X(t)=(A+A^2t+\frac{t^2}{2!}A^3+\cdots+\frac{t^n}{n!}A^{n+1}+\cdots)K \tag{4-9}$$

比较(4-8)及(4-9)两式，即可得出

$$\frac{d}{dt}(e^{At}K)=Ae^{At}K$$

这就证明了$e^{At}K$是方程(4-7)的解。

把解方程(4-7)的问题，转化为矩阵函数e^{At}的计算问题(e^{At}的几种计算方法问题将在下节中讲述)。于是满足初始条件$X(0)=X_0$的特解为

$$X(t)=e^{At}X_0 \tag{4-10}$$

二、常系数线性非齐次方程

这是一种在受外力作用的动力系统中常见的方程。考虑矩阵微分方程

$$\frac{d}{dt}X(t)=AX(t)+Bu(t) \tag{4-11}$$

$$X(0)=X_0$$

其中X是n维向量，u是r维向量，$A=[a_{ij}]_{n\times n}$，$B=[b_{ij}]_{n\times r}$均为常数矩阵。方程(4-11)改写为

$$\frac{d}{dt}X(t)-AX(t)=Bu(t)$$

两边左乘e^{-At}，得到

$$e^{-At}\frac{d}{dt}X(t)-e^{-At}AX(t)=e^{-At}Bu(t)$$

即 $\frac{d}{dt}[e^{-At}X(t)]=e^{-At}Bu(t)$

对上式积分并由$X(0)=X_0$可得

$$e^{-At}X(t)=X_0+\int_0^t e^{-At}Bu(t)dt \qquad (4\text{-}12)$$

这就是方程（4-11）的解。

即 $X(t)=e^{At}X_0+\int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$

第三节　e^{At}的计算方法

〔**方法1**〕　利用Jordan标准形计算

设：$A=[a_{ij}]_{n\times n}$，将A化为Jordan标准形。若A的特征值不同，即

$|\lambda I-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，其中$\lambda_i\neq\lambda_j\ (i\neq j)$

由此得到

$$A\sim J=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

$A=PJP^{-1}$，其中$P=[p_{ij}]_{n\times n}$

根据上章公式(3-52)

$$e^{Jt}=\begin{pmatrix}e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t}\end{pmatrix}$$

于是

$$e^{At}=Pe^{Jt}P^{-1}$$

$$=P\begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t} \end{bmatrix}P^{-1} \tag{4-13}$$

必须指出，若A含有重特征值，则e^{At}就需用稍为复杂的形式，即需借助上章公式(3-54)来写出。

〔**方法2**〕 利用Laplace变换法

对方程 $\frac{d}{dt}X(t)=AX(t)$，$X(0)=X_0$

施行Laplace变换，得

$$sF(s)-X(0)=AF(s)$$

或 $(sI-A)F(s)=X_0$

用矩阵$(sI-A)^{-1}$左乘上式两边，有

$$F(s)=(sI-A)^{-1}X_0 \tag{4-14}$$

取逆变换即得到

$$X(t)=L^{-1}[(sI-A)^{-1}]X_0$$

由上节知方程（4-7）满足初始条件的特解为：

$$X(t)=e^{At}X_0$$

于是有 $e^{At}=L^{-1}[(sI-A)^{-1}]$ (4-15)

即 $L(e^{At})=(sI-A)^{-1}$

根据逆矩阵运算公式，有

$$(sI-A)^{-1}=\frac{(sI-A)}{|sI-A|} \tag{4-16}$$

这在较简单情况下才可直接得出这一运算结果。

如果$(sI-A)^{-1}$具有较为复杂的形式，例如A的阶数大于

3时，计算是比较复杂的，此时可用以下方法求得：

设 $P=(sI-A)^{-1}$

由于 $(sI-A)P=(sI-A)(sI-A)^{-1}=I$

于是 $sP=AP+I$

两边左乘（$sI+A$），即得到

$$s^2P=A^2P+A+sI$$

两边再左乘（$sI+A$），又得到

$$s^3P=A^3P+A^2+sA+s^2I$$

重复使用这一过程，即得出一组方程：

$$P=P$$

$$sP=AP+I$$

$$s^2P=A^2P+A+sI$$

$$s^3P=A^3P+A^2+sA+s^2I$$

………………………………

$$s^pP=A^pP+A^{p-1}+sA^{p-2}+\cdots+s^{p-2}A+s^{p-1}I$$

其中：p是A的最小多项式的次数。

设A的最小多项式为：

$$\varphi_m(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_p\lambda^p,$$

则有 $a_0I+a_1A+a_2A^2+\cdots+a_pA^p=0$

其中： $a_p=1$

将上方程组逐个乘以a_0，a_1，…，a_p后，两边相加，并利用最小多项式为零化多项式这一关系，可得到

$$\left(\sum_{i=0}^{p}a_is^i\right)P=\sum_{i=0}^{p}a_iA^iP+\sum_{i=1}^{p}a_iA^{i-1}+s\sum_{i=2}^{p}a_iA^{i-2}$$

$$+\cdots+s^{p-2}\sum_{i=p-1}^{p}a_iA^{i-p+1}+s^{p-1}a_pI$$

$$= \sum_{j=0}^{p-1} s^j \sum_{i=1+j}^{p} a_i A^{i-j-1}$$

所以 $$(sI-A)^{-1}=P=\frac{\sum_{j=0}^{p-1} s^j \sum_{i=1+j}^{p} a_i A^{i-j-1}}{\sum_{i=0}^{p} a_i s^i} \tag{4-17}$$

如果A的最小多项式与特征多项式全同，即$p=n$，则方程(4-17)即成为：

$$(sI-A)^{-1}=\frac{\sum_{j=0}^{n-1} s^j \sum_{i=1+j}^{n} a_i A^{i-j-1}}{|sI-A|} \tag{4-18}$$

其中 $|sI-A|=\sum_{i=0}^{n} a_i s^i, \qquad a_n=1$

〔**方法3**〕 利用矩阵函数的定义或矩阵函数的基本公式(3-55)均可计算e^{At}。

〔**方法4**〕 e^{At}的数值计算法（略去）。

〔**例1**〕 求矩阵微分方程

$$\frac{d}{dt}X=AX \tag{4-19}$$

满足初始条件 $X(0)=X_0$的特解。

其中：

$$X=\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}; \quad A=\begin{pmatrix} -7 & -7 & 5 \\ -8 & -8 & -5 \\ 0 & -5 & 0 \end{pmatrix};$$

$$X(0)=X_0=\begin{pmatrix}3\\-2\\1\end{pmatrix}$$

解　把A化成Jordan标准形

$$|\lambda I-A|=\begin{vmatrix}\lambda+\eta & \eta & -5\\ 8 & \lambda+8 & 5\\ 0 & 5 & \lambda\end{vmatrix}=(\lambda-5)(\lambda+5)(\lambda+15)$$

因此，特征值$\lambda_1=5$，$\lambda_2=-5$，$\lambda_3=-15$

$$A\sim J=\begin{pmatrix}5 & & \\ & -5 & \\ & & -15\end{pmatrix}$$

由　$A=PJP^{-1}$得　$AP=PJ$，设$P=[\eta_1,\ \eta_2,\ \eta_3]$

则有

$$A[\eta_1,\eta_2,\eta_3,]=[\eta_1,\eta_2,\ \eta_3]\begin{pmatrix}5 & & \\ & -5 & \\ & & -15\end{pmatrix}$$

比较上式两边，得

$$\eta_1=\begin{pmatrix}1\\-1\\1\end{pmatrix};\quad \eta_2=\begin{pmatrix}1\\-1\\-1\end{pmatrix};\quad \eta_3=\begin{pmatrix}2\\3\\1\end{pmatrix}$$

故满秩矩阵 $P=\begin{pmatrix}1 & 1 & 2\\ -1 & -1 & 3\\ 1 & -1 & 1\end{pmatrix}$

再由上章公式(3-52)可得

$$e^{At}=P\begin{pmatrix}e^{5t} & & \\ & e^{-5t} & \\ & & e^{-15t}\end{pmatrix}P^{-1}$$

$$=\begin{pmatrix}1 & 1 & 2\\-1 & -1 & 3\\1 & -1 & 1\end{pmatrix}\begin{pmatrix}e^{5t} & & \\ & e^{-5t} & \\ & & e^{-15t}\end{pmatrix}\begin{pmatrix}1 & 1 & 2\\-1 & -1 & 3\\1 & -1 & 1\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}e^{5t} & e^{-5t} & 2e^{-15t}\\-e^{5t} & -e^{-5t} & 3e^{-15t}\\e^{5t} & -e^{-5t} & e^{-15t}\end{pmatrix}\begin{pmatrix}2 & -3 & 5\\4 & -1 & -5\\2 & 2 & 0\end{pmatrix}\times\frac{1}{10}$$

$$=\frac{1}{10}\begin{pmatrix}2e^{5t}+4e^{-5t}+4e^{-15t} & -3e^{5t}-e^{-5t}+4e^{-15t} & 5e^{5t}-5e^{-5t}\\-2e^{5t}-4e^{-5t}+6e^{-15t} & 3e^{5t}+e^{-5t}+6e^{-15t} & -5e^{5t}+5e^{-5t}\\2e^{5t}-4e^{-5t}+2e^{-15t} & -3e^{5t}+e^{-5t}+2e^{-15t} & 5e^{5t}+5e^{-5t}\end{pmatrix}$$

根据前面第二节第一部分知，方程(4-19)满足初始条件的特解为：

$$X(t)=e^{At}X_0$$

$$=\frac{1}{10}\begin{pmatrix}17e^{5t}+9e^{-5t}+4e^{-15t}\\-17e^{5t}-9e^{-5t}+6e^{-15t}\\17e^{5t}-9e^{-5t}+2e^{-15t}\end{pmatrix}$$

即 $x_1(t)=\frac{17}{10}e^{5t}+\frac{9}{10}e^{-5t}+\frac{2}{5}e^{-15t}$

$$x_2(t)=-\frac{17}{10}e^{5t}-\frac{9}{10}e^{-5t}+\frac{3}{5}e^{-15t}$$

$$x_3(t)=\frac{17}{10}e^{5t}-\frac{9}{10}e^{-5t}+\frac{1}{5}e^{-15t}$$

用方法 2 亦可求得相同答案，读者可自行解算。

利用矩阵函数的基本公式(3-55)求出e^{At}，也可得到相同的结果。

因矩阵A的最小多项式就是A的特征多项式（第二章第三节中定理4的推论）。

$$\varphi_m(\lambda)=(\lambda-5)(\lambda+5)(\lambda+15)$$

设 $f(\lambda)=e^{\lambda t}$，由（3-55）式知

$$f(A)=Z_1(A)f(5)+Z_2(A)f(-5)+Z_3(A)f(-15)$$

可取$f_1(\lambda)=(\lambda+5)(\lambda+15)$，得 $f_1(5)=200$

$f_1(-5)=f_1(-15)=0$

$f_2(\lambda)=(\lambda-5)(\lambda+15)$，得 $f_2(-5)=-100$，

$f_2(5)=f_2(-15)=0$

$f_3(\lambda)=(\lambda-5)(\lambda+5)$，得 $f_3(-15)=200$

$f_3(5)=f_3(-5)=0$

于是，得

$$Z_1(A)=\frac{1}{200}(A+5I)(A+15I)=\frac{1}{200}(A^2+20A+75I)$$

$$Z_2(A)=-\frac{1}{100}(A-5I)(A+15I)$$

$$=-\frac{1}{100}(A^2+10A-75I)$$

$$Z_3(A)=\frac{1}{200}(A-5I)(A+5I)=\frac{1}{200}(A^2-25I)$$

$$f(A)=\frac{1}{200}(A^2+20A+75I)f(5)-\frac{1}{100}(A^2+10A$$

$$-75I)f(-5)+\frac{1}{200}(A^2-25I)f(-15)$$

故

$$f(A)=e^{At}=\frac{1}{200}(A^2+2A+75I)e^{5t}-\frac{1}{100}(A^2+10A$$

$$-75I)e^{-5t}+\frac{1}{200}(A^2-25I)e^{-15t}$$

用这种方法求e^{At}较为方便，因为不需求变换矩阵P，只计算A^2即可求得。

〔**例2**〕 求矩阵微分方程

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix}=\begin{bmatrix}-6 & 1 & 0\\ -11 & 0 & 1\\ -6 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}+\begin{bmatrix}2\\ 6\\ 2\end{bmatrix}u \tag{4-20}$$

满足初始条件

$$X(0)=\begin{bmatrix}x_1(0)\\ x_2(0)\\ x_3(0)\end{bmatrix}$$

的特解。其中$u=u(t)$为单位阶跃函数，即

$$u(t)=\begin{cases}1 & t\geqslant 0\\ 0 & t<0\end{cases}$$

解 给定方程可视为

$$\dot{X}=AX(t)+Bu(t)$$

将

$$A=\begin{bmatrix}-6 & 1 & 0\\ -11 & 0 & 1\\ -6 & 0 & 0\end{bmatrix}\quad \text{化为Jordan标准形}$$

$$|\lambda I-A|=\begin{vmatrix}\lambda+6 & -1 & 0\\ 11 & \lambda & -1\\ 6 & 0 & \lambda\end{vmatrix}=(\lambda+1)(\lambda+2)(\lambda+3)$$

所以特征值为：$\lambda_1=-1$，$\lambda_2=-2$，$\lambda_3=-3$

于是 $A\sim J=\begin{bmatrix}-1 & & \\ & -2 & \\ & & -3\end{bmatrix}$

由$A=PJP^{-1}$，得$AP=PJ$，设$P=[\eta_1, \eta_2, \eta_3]$

则有：$A[\eta_1, \eta_2, \eta_3]=[\eta_1, \eta_2, \eta_3]\begin{bmatrix}-1 & & \\ & -2 & \\ & & -3\end{bmatrix}$

比较上式两边，可得

$$\eta_1=\begin{bmatrix}1\\5\\6\end{bmatrix},\quad \eta_2=\begin{bmatrix}1\\4\\3\end{bmatrix},\quad \eta_3=\begin{bmatrix}1\\3\\2\end{bmatrix}$$

故得满秩矩阵

$$P=\begin{bmatrix}1 & 1 & 1\\5 & 4 & 3\\6 & 3 & 2\end{bmatrix}$$

又由上章公式(3-52)可得

$$e^{At}=P\begin{bmatrix}e^{-t} & & \\ & e^{-2t} & \\ & & e^{-3t}\end{bmatrix}P^{-1}$$

$$=\begin{bmatrix}1 & 1 & 1\\5 & 4 & 3\\6 & 3 & 2\end{bmatrix}\begin{bmatrix}e^{-t} & & \\ & e^{-2t} & \\ & & e^{-3t}\end{bmatrix}\begin{bmatrix}-1 & 1 & -1\\8 & -4 & 2\\-9 & 3 & -1\end{bmatrix}\left(-\frac{1}{2}\right)$$

$$=\begin{pmatrix} e^{-t} & e^{-2t} & e^{-3t} \\ 5e^{-t} & 4e^{-2t} & 3e^{-3t} \\ 6e^{-t} & 3e^{-2t} & 2e^{-3t} \end{pmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 2 & -1 \\ \frac{9}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$$

根据第二节第二部分知，方程(4-20)满足初始条件的特解为：

$$X(t)=e^{At}X_0+\int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

$$=P\begin{bmatrix} e^{-t} & & \\ & e^{-2t} & \\ & & e^{-3t} \end{bmatrix}P^{-1}\begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix}$$

$$+P\int_0^t \begin{bmatrix} e^{-(t-\tau)} & & \\ & e^{-2(t-\tau)} & \\ & & e^{-3(t-\tau)} \end{bmatrix}P^{-1}Bu(\tau)d\tau$$

$$=\begin{bmatrix} e^{-t} & e^{-2t} & e^{-3t} \\ 5e^{-t} & 4e^{-2t} & 3e^{-3t} \\ 6e^{-t} & 3e^{-2t} & 2e^{-3t} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 2 & -1 \\ \frac{9}{2} & -\frac{3}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix}$$

$$+\begin{bmatrix} 1 & 1 & 1 \\ 5 & 4 & 3 \\ 6 & 3 & 2 \end{bmatrix}\int_0^t \begin{bmatrix} e^{-(t-\tau)} & & \\ & e^{-2(t-\tau)} & \\ & & e^{-3(t-\tau)} \end{bmatrix}$$

$$\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 2 & -1 \\ \frac{9}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix} d\tau$$

$$= \begin{pmatrix} e^{-t} & e^{-2t} & e^{-3t} \\ 5e^{-t} & 4e^{-2t} & 3e^{-3t} \\ 6e^{-t} & 3e^{-2t} & 2e^{-3t} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 2 & -1 \\ \frac{9}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{pmatrix}$$

$$+ \begin{pmatrix} 1 & 1 & 1 \\ 5 & 4 & 3 \\ 6 & 3 & 2 \end{pmatrix} \int_0^t \begin{pmatrix} -e^{-(t-\tau)} \\ 2e^{-2(t-\tau)} \\ e^{-3(t-\tau)} \end{pmatrix} d\tau$$

$$= \begin{pmatrix} e^{-t} & e^{-2t} & e^{-3t} \\ 5e^{-t} & 4e^{-2t} & 3e^{-3t} \\ 6e^{-t} & 3e^{-2t} & 2e^{-3t} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -4 & 2 & -1 \\ \frac{9}{2} & -\frac{3}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{pmatrix}$$

$$+ \begin{pmatrix} 1 & 1 & 1 \\ 5 & 4 & 3 \\ 6 & 3 & 2 \end{pmatrix} \begin{pmatrix} e^{-t}-1 \\ 1-e^{-2t} \\ \frac{1}{3}\left(1-e^{-3t}\right) \end{pmatrix}$$

因此解为：

$$
X(t)=\begin{pmatrix}x_1(t)\\x_2(t)\\x_3(t)\end{pmatrix}=\begin{pmatrix}e^{-t} & e^{-2t} & e^{-3t}\\5e^{-t} & 4e^{-2t} & 3e^{-3t}\\6e^{-t} & 3e^{-2t} & 2e^{-3t}\end{pmatrix}\begin{pmatrix}\frac{1}{2} & -\frac{1}{2} & \frac{1}{2}\\-4 & 2 & -1\\\frac{9}{2} & -\frac{3}{2} & \frac{1}{2}\end{pmatrix}
$$

$$
\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+\begin{pmatrix}1 & 1 & 1\\5 & 4 & 3\\6 & 3 & 2\end{pmatrix}\begin{pmatrix}e^{-t}-1\\1-e^{-2t}\\\frac{1}{3}(1-e^{-3t})\end{pmatrix}
$$

第四节　状态转移矩阵

一、状态转移矩阵的性质

对于矩阵微分方程

$$
\frac{d}{dt}X(t)=AX(t) \tag{4-21}
$$

在第二节中已讨论了满足初始条件

$$
X(0)=X_0
$$

的解是：

$$
X(t)=e^{At}X_0=e^{At}X(0) \tag{4-22}
$$

这样，方程(4-21)的求解问题就转化为求矩阵函数 e^{At}的问题。在工程问题中常把矩阵函数e^{At}记为$\Phi(t)$，即$\Phi(t)=e^{At}$于是(4-22)式就可表示为：

$$
X(t)=\Phi(t)X(0) \tag{4-23}
$$

可以看出，方程（4-21）的解，只不过是初始状态的一个简单转移，把$t=0$时刻的初始状态$X(0)=X_0$转移到 t 时刻的状

态$X(t)$。因此，把矩阵$\Phi(t)$称为**状态转移矩阵**（或称基本矩阵）。这个矩阵包含了所描述系统的全部自由运动信息，在工程研究中广为应用。

状态转移矩阵$\Phi(t)$具有以下几个主要性质：

性质1　$\Phi(0)=I$，其中I为单位矩阵。

证明　由第三章矩阵分析理论知

$$e^{At}=I+At+\frac{1}{2!}A^2t^2+\cdots+\frac{1}{n!}A^nt^n+\cdots \tag{4-24}$$

在（4-24）式中，代入$t=0$，得

$$\Phi(0)=e^{A\cdot 0}=I$$

性质2　$\Phi^{-1}(t)=\Phi(-t)$，或$\Phi(t)=[\Phi(-t)]^{-1}$

证明　首先证明

$$e^{A(t+s)}=e^{At}e^{As}$$

由于 $e^{At}e^{As}=\left(\sum_{n=0}^{\infty}\frac{A^nt^n}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{A^ns^n}{n!}\right)$

$$=\sum_{n=0}^{\infty}A^n\left(\sum_{i=0}^{n}\frac{n!}{i!}\frac{t^is^{n-i}}{(n-i)!}\right)\frac{1}{n!}$$

$$=\sum_{n=0}^{\infty}A^n\frac{(t+s)^n}{n!}=e^{A(t+s)}$$

特别地，若$s=-t$，有

$$e^{At}e^{-At}=e^{A(t-t)}=I \tag{4-25}$$

由此可知，e^{At}的逆为e^{-At}

由于e^{At}的逆总是存在的，所以e^{At}是非奇异矩阵。于是(4-25)式可写为：

$$\Phi(t)\Phi(-t)=\Phi(-t)\Phi(t)=I$$

故有

$$\Phi^{-1}(t)=\Phi(-t)$$

或 $$\Phi(t)=[\Phi(-t)]^{-1}$$ 证毕。

性质3 $\Phi(t_1+t_2)=\Phi(t_1)\Phi(t_2)=\Phi(t_2)\Phi(t_1)$

证明 由于

$$\Phi(t_1+t_2)=e^{A(t_1+t_2)}=e^{At_1}e^{At_2}=e^{At_2}e^{At_1}$$

故有

$$\Phi(t_1+t_2)=\Phi(t_1)\Phi(t_2)=\Phi(t_2)\Phi(t_1)$$

性质4 $[\Phi(t)]^n=\Phi(nt)$

证明 因为

$$[\Phi(t)]^n=\underbrace{\Phi(t)\Phi(t)\cdots\Phi(t)}_{n个}=\underbrace{e^{At}e^{At}\cdots e^{At}}_{n个}=e^{Ant}$$

故有 $$[\Phi(t)]^n=\Phi(nt)$$

性质5 $$\Phi(t_2-t_1)\Phi(t_1-t_0)=\Phi(t_1-t_0)\Phi(t_2-t_1)=\Phi(t_2-t_0)$$

证明 由于

$$e^{A(t_2-t_1)}e^{A(t_1-t_0)}=e^{A(t_2-t_0)}=e^{A(t_1-t_0+t_2-t_1)}=e^{A(t_1-t_0)}e^{A(t_2-t_1)}$$

所以

$$\Phi(t_2-t_1)\Phi(t_1-t_0)=\Phi(t_2-t_0)=\Phi(t_1-t_0)\Phi(t_2-t_1)$$

最后，关于状态转移矩阵$\Phi(t)=e^{At}$的计算在本章第三节中已进行讲述，今再通过下列讲述，阐明利用Laplace变换来计算$\Phi(t)$。

〔**例1**〕 求齐次状态方程

$$\dot{X}=AX \tag{4-26}$$

的状态转移矩阵$\Phi(t)$和状态转移矩阵的逆$\Phi^{-1}(t)$。

其中 $A=\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$， $X=\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$

解 由(4-15)式可知，状态转移矩阵$\Phi(t)$可确定为

$$\Phi(t)=e^{At}=L^{-1}[(sI-A)^{-1}]$$

因为 $$A=\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$$

于是 $$sI-A=\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}-\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}=\begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix}$$

而 $$(sI-A)^{-1}=\frac{1}{(s+1)(s+2)}\begin{bmatrix} s+3 & 1 \\ -2 & s \end{bmatrix}$$

$$=\begin{bmatrix} \dfrac{s+3}{(s+1)(s+2)} & \dfrac{1}{(s+1)(s+2)} \\ \dfrac{-2}{(s+1)(s+2)} & \dfrac{s}{(s+1)(s+2)} \end{bmatrix}$$

所以 $$\Phi(t)=e^{At}=L^{-1}[(sI-A)^{-1}]$$

$$=\begin{bmatrix} 2e^{-t}-e^{-2t} & e^{-t}-e^{-2t} \\ -2e^{-t}+2e^{-2t} & -e^{-t}+2e^{-2t} \end{bmatrix}$$③

由性质2有 $\Phi^{-1}(t)=\Phi(-t)$

故状态转移矩阵的逆应为：

$$\Phi^{-1}(t)=\Phi(-t)=\begin{bmatrix} 2e^{t}-e^{2t} & e^{t}-e^{2t} \\ -2e^{t}+2e^{2t} & -e^{t}+2e^{2t} \end{bmatrix}$$

二、利用状态转移矩阵求解非齐次矩阵微分方程

---

③矩阵的逆Laplace变换，为所有元素的逆Laplace变换所构成的矩阵。

利用状态转移矩阵及其性质，可以很方便地求得非齐次矩阵微分方程

$$\dot{X}(t)=AX(t)+Bu(t) \tag{4-27}$$

满足 $X(0)=X_0$

的解。

其中：$X(t)$表示n维向量；$u(t)$表示r维向量；

$A=[a_{ij}]_{n\times n}$ ； $B=[b_{ij}]_{n\times r}$ 均为常数矩阵。由第二节第二部分知，方程（4-27）的解为：

$$e^{-At}X(t)=X(0)+\int_0^t e^{-A\tau}Bu(\tau)d\tau$$

或

$$X(t)=e^{At}X_0+\int_0^t e^{A(t-\tau)}Bu(\tau)d\tau \tag{4-28}$$

利用状态转移矩阵$\Phi(t)=e^{At}$，上式可改写为：

$$X(t)=\Phi(t)X_0+\int_0^t \Phi(t-\tau)Bu(\tau)d\tau \tag{4-29}$$

解$X(t)$中包含有初始状态的转移项和由输入向量所产生的项。

〔**例2**〕 求非齐次状态方程

$$\dot{X}(t)=AX(t)+Bu(t) \tag{4-30}$$

满足 $X(0)=X_0$

的解。

其中

$$A=\begin{bmatrix}0 & 1\\ -2 & -3\end{bmatrix};\quad B=\begin{bmatrix}0\\ 1\end{bmatrix};\quad X(t)=\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix};$$

$$u(t)=\begin{cases}1 & t\geqslant 0\\ 0 & t<0\end{cases}$$

解 系数矩阵A已在例1中求得其状态转移矩阵$\Phi(t)=e^{At}$为：

$$\Phi(t)=e^{At}=\begin{pmatrix} 2e^{-t}-e^{-2t} & e^{-t}-e^{-2t} \\ -2e^{-t}+2e^{-2t} & -e^{-t}+2e^{-2t} \end{pmatrix}$$

自由项矩阵为 $B=\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

由（4-29）式可得方程（4-30）的解为

$$X(t)=\Phi(t)X(0)$$

$$+\int_0^t \begin{pmatrix} 2e^{-(t-\tau)}-e^{-2(t-\tau)} & e^{-(t-\tau)}-e^{-2(t-\tau)} \\ -2e^{-(t-\tau)}+2e^{-2(t-\tau)} & -e^{-(t-\tau)}+2e^{-2(t-\tau)} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix}d\tau$$

或 $$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=\begin{pmatrix} 2e^{-t}-e^{-2t} & e^{-t}-e^{-2t} \\ -2e^{-t}+2e^{-2t} & -e^{-t}+2e^{-2t} \end{pmatrix}\begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}$$

$$+\begin{pmatrix} \frac{1}{2}-e^{-t}+\frac{1}{2}e^{-2t} \\ e^{-t}-e^{-2t} \end{pmatrix}$$

若初始状态是$X(0)=0$，则上式即简化为：

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=\begin{pmatrix} \frac{1}{2}-e^{-t}+\frac{1}{2}e^{-2t} \\ e^{-t}-e^{-2t} \end{pmatrix}$$

第五节　变系数线性矩阵微分方程（线性时变系统）

我们可以利用状态转移矩阵来求变系数矩阵微分方程的解。只要将状态转移矩阵$\Phi(t)$改为$\Phi(t,\ t_0)$，在第四节中讨论的大部分结论都可用来讨论变系数矩阵微分方程的解。但

是应注意到，时变系统的状态转移矩阵既与t有关，也与t_0有关，而与t和t_0之差无关。因此，虽然在有些情况下t_0可取为零，但不是总取为零；此外对时变系统来说，状态转移矩阵通常不能用指数矩阵函数e^{At}给出。

一、齐次矩阵微分方程的解

设变系数矩阵微分方程为

$$\dot{X}(t)=A(t)X(t) \tag{4-31}$$

满足 $X(t_0)=X_0$

其中：$X(t)$为n维向量； $A(t)=[a_{ij}(t)]_{n\times n}$ ，$a_{ij}(t)$ 在时间区间$[t_0,\ t_1]$上是t的分段连续函数。

则方程（4-31)的解为

$$X(t)=\Phi(t,\ t_0)X(t_0) \tag{4-32}$$

其中：$\Phi(t,\ t_0)$ 满足矩阵微分方程

$$\dot{\Phi}(t,\ t_0)=A(t)\Phi(t,\ t_0) \tag{4-33}$$

及 $\Phi(t_0,t_0)=I$，且$\Phi(t,t_0)=[\varphi_{ij}(t,t_0)]_{n\times n}$是非奇异矩阵。

由此，$X(t_0)=\Phi(t_0,\ t_0)X(t_0)=IX(t_0)=X_0$

且 $$\dot{X}(t)=\frac{d}{dt}[\Phi(t,\ t_0)X(t_0)]=\dot{\Phi}(t,\ t_0)X(t_0)$$

$$=A(t)\Phi(t,\ t_0)X(t_0)=A(t)X(t)$$

从而可见，方程（4-31）的解只不过是初始状态的转移。因此，矩阵$\Phi(t,\ t_0)$ 就是方程（4-31）所描述的时变系统的状态转移矩阵。具有以下一些性质：

性质1 $\Phi(t_2,\ t_1)\Phi(t_1,\ t_0)=\Phi(t_2,\ t_0)$

证明 由于 $X(t_1)=\Phi(t_1,\ t_0)X(t_0)$

$X(t_2)=\Phi(t_2,\ t_1)X(t_1)$

于是

$$X(t_2)=\Phi(t_2,\ t_1)\Phi(t_1,\ t_0)X(t_0)=\Phi(t_2,\ t_0)X(t_0)$$

故有

$$\Phi(t_2, t_1)\Phi(t_1, t_0)=\Phi(t_2, t_0)$$

性质2 $\Phi(t_1, t_0)=\Phi^{-1}(t_0, t_1)$

证明 由性质1可得

$$\Phi(t_1, t_0)=\Phi^{-1}(t_2, t_1)\Phi(t_2, t_0)$$

取$t_2=t_0$代入上式，则得：

$$\Phi(t_1, t_0)=\Phi^{-1}(t_0, t_1)\Phi(t_0, t_0)=\Phi^{-1}(t_0, t_1)。$$

应注意：时变系统的状态转移矩阵

$$\Phi(t, t_0)=e^{\int_{t_0}^{t} A(\tau)d\tau}$$

只在$A(t)$和$\int_{t_0}^{t} A(\tau)d\tau$相乘时，对所有$t$都可以交换的情况下才成立④。若$A(t)$是一个常数矩阵或对角线矩阵，则

---

④对于方程 $\frac{dX}{dt}=A(t)X(t)$，其解为：$X(t)=\Phi(t, t_0)X(t_0)$

其中$\Phi(t, t_0)=e^{\int_{t_0}^{t} A(\tau)d\tau}=I+\int_{t_0}^{t}A(\tau)d\tau+\frac{1}{2!}\left[\int_{t_0}^{t}A(\tau)d\tau\right]^2$

$$+\frac{1}{3!}\left[\int_{t_0}^{t}A(\tau)d\tau\right]^3+\cdots$$

将此解 $X(t)=\phi(t, t_0)X(t_0)$代入原方程左、右两边分别得：

$$\frac{dX}{dt}=\left\{A(t)+\left(\int_{t_0}^{t}A(\tau)d\tau\right)A(t)\right.$$

$$\left.+\frac{1}{2!}\left(\int_{t_0}^{t}A(\tau)d\tau\right)^2A(\tau)\quad+\cdots\right\}X(t_0)$$

$$A(t)X(t)=\left\{A(t)+A(t)\left(\int_{t_0}^{t}A(\tau)d(\tau)\right)\right.$$

$$\left.+\frac{1}{2!}A(t)\left(\int_{t_0}^{t}A(\tau(d\tau\right)^2+\cdots\right\}X(t_0)$$

显然，只在$A(t)$和$\int_{t_0}^{t}A(\tau)d\tau$可交换时，以上两式才能相等。因此，只有在这种情况下$X(t)=\phi(t, t_0)X(t_0)$才是方程$\frac{dX}{dt}=A(t)X(t)$的解。

$A(t)$与$\int_{t_0}^{t} A(\tau)d\tau$是可换的。对一般时变矩阵$A(t)$来说，常常是对$\Phi(t, t_0)$作数值积分，即将$\Phi(t, t_0)$展成级数形式：

$$\Phi(t, t_0) = I + \int_{t_0}^{t} A(\tau)d\tau + \int_{t_0}^{t} A(\tau_1)\left[\int_{t_0}^{\tau_1} A(\tau_2)d\tau_2\right] d\tau_1 + \int_{t_0}^{t} A(\tau_1)\left[\int_{t_0}^{\tau_1} A(\tau_2)\left(\int_{t_0}^{\tau_2} A(\tau_3)d\tau_3\right)d\tau_2\right]d\tau_1 + \cdots \tag{4-34}$$

通常不能用封闭形式给出$\Phi(t, t_0)$。

〔**例**〕 求变系数矩阵微分方程

$$\dot{X} = A(t)X$$

的状态转移矩阵$\Phi(t, 0)$。

其中 $A(t) = \begin{bmatrix} 0 & 1 \\ 0 & t \end{bmatrix}$； $X = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$

解 根据公式（4-34）

$$\Phi(t, 0) = I + \int_0^t A(\tau)d\tau + \int_0^t A(\tau_1)\left(\int_0^{\tau_1} A(\tau_2)d\tau_2\right)d\tau_1 + \cdots$$

因为

$$\int_0^t A(\tau)d\tau = \int_0^t \begin{bmatrix} 0 & 1 \\ 0 & \tau \end{bmatrix} d\tau = \begin{bmatrix} 0 & t \\ 0 & \frac{t^2}{2} \end{bmatrix} \int_0^t \begin{bmatrix} 0 & 1 \\ 0 & \tau_1 \end{bmatrix}$$

$$\left(\int_0^{\tau_1} \begin{bmatrix} 0 & 1 \\ 0 & \tau_2 \end{bmatrix} d\tau_2\right) d\tau_1 = \int_0^t \begin{bmatrix} 0 & 1 \\ 0 & \tau_1 \end{bmatrix} \begin{bmatrix} 0 & \tau_1 \\ 0 & \frac{\tau_1^2}{2} \end{bmatrix} d\tau_1$$

$$=\begin{bmatrix} 0 & \frac{t^3}{6} \\ 0 & \frac{t^4}{8} \end{bmatrix}$$

故得 $$\Phi(t,\ 0)=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}+\begin{bmatrix} 0 & t \\ 0 & \frac{t^2}{2} \end{bmatrix}+\begin{bmatrix} 0 & \frac{t^3}{6} \\ 0 & \frac{t^4}{8} \end{bmatrix}+\cdots$$

$$=\begin{bmatrix} 1 & t+\frac{t^3}{6}+\cdots \\ 0 & 1+\frac{t^2}{2}+\frac{t^4}{8}+\cdots \end{bmatrix}$$

二、非齐次矩阵微分方程的解

现在来讨论以下状态方程的求解。

$$\dot{X}(t)=A(t)X(t)+B(t)u(t) \tag{4-35}$$

$$X(t_0)=X_0$$

其中：X表示n维向量；u表示r维向量。

$A(t)=[a_{ij}(t)]_{n\times n}$；$B(t)=[b_{ij}(t)]_{n\times r}$

$a_{ij}(t)$，$b_{ij}(t)$在区间$[t_0,\ t_1]$内均为分段连续函数。

设方程（4-35）的解为 $X(t)=\Phi(t,\ t_0)Y(t)$

$\Phi(t,\ t_0)$是满足下面方程的唯一矩阵。

$\dot{\Phi}(t,\ t_0)=A(t)\Phi(t,\ t_0)$及$\Phi(t_0,\ t_0)=I$

则 $$\dot{X}(t)=\frac{d}{dt}\left[\Phi(t,\ t_0)Y(t)\right]$$

$$=\dot{\Phi}(t,\ t_0)Y(t)+\Phi(t,t_0)\dot{Y}(t)$$

$$=A(t)\Phi(t,\ t_0)Y(t)+\Phi(t,\ t_0)\dot{Y}(t)$$

又 $\dot{X}(t)=A(t)\Phi(t,\ t_0)Y(t)+B(t)u(t)$

所以必有

$$\Phi(t,\ t_0)\dot{Y}(t)=B(t)u(t)$$

即 $\dot{Y}(t)=\Phi^{-1}(t,\ t_0)B(t)u(t)$

于是 $Y(t)=\int_{t_0}^{t}\Phi^{-1}(\tau,\ t_0)B(\tau)u(\tau)\mathrm{d}\tau+Y(t_0)$

由于 $Y(t_0)=\Phi^{-1}(t_0,\ t_0)X(t_0)=X(t_0)$

因此，方程（4-35）的解为

$$X(t)=\Phi(t,\ t_0)X(t_0)+\Phi(t,\ t_0)\int_{t_0}^{t}\Phi^{-1}(\tau,t_0)B(\tau)u(\tau)\mathrm{d}\tau$$

$$=\Phi(t,t_0)X(t_0)+\int_{t_0}^{t}\Phi(t,\ \tau)B(\tau)u(\tau)\mathrm{d}\tau \qquad (4\text{-}36)$$

这就得到了非齐次矩阵微分方程的解，是其相应的齐次微分方程的解$\Phi(t,t_0)X_0$与一个特解$X_p(t)$的和。

特解$X_p(t)$为：

$$X_p(t)=\Phi(t,t_0)\int_{t_0}^{t}\Phi^{-1}(\tau,t_0)B(\tau)u(\tau)\mathrm{d}\tau$$

$$=\int_{t_0}^{t}\Phi(t,\tau)B(\tau)u(\tau)\mathrm{d}\tau$$

公式（4-36）需要用电子计算机来计算。

第六节　矩阵Riccati方程

在最优控制规律研究中，经常需要求解一种Riccati类型的方程。这里我们介绍一种应用矩阵分析理论来求解矩阵Riccati方程的方法。为此，我们考虑如下的矩阵Riccati方程：

$$\dot{P}(t)=A(t)P(t)+P(t)B(t)+P(t)C(t)P(t)+D(t) \tag{4-37}$$

满足 $P(t_0)=E$

其中$P(t)$是$n\times n$阶未知矩阵；$A(t)$，$B(t)$，$C(t)$，$D(t)$都是$n\times n$阶已知矩阵，且其元素在区间$[t_0, t_1]$上是绝对可积的。

下面的解法，是把方程（4-37）化成另一个矩阵微分方程

$$\dot{W}(t)=G(t)W(t) \tag{4-38}$$

$$W(t_0)=\begin{pmatrix} I \\ E \end{pmatrix}$$

其中：
$$W(t)=\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix},\ G(t)=\begin{pmatrix} -B & -C \\ D & A \end{pmatrix}$$

来求解。若能在$[t_0, t_1]$中解得矩阵$Y(t)$和可逆阵$X(t)$，则Riccati方程（4-37）的解，就可由

$$P(t)=Y(t)X^{-1}(t) \qquad t\in[t_0, t_1]$$

给出。这可证明如下：

1° YX^{-1}满足方程（4-37）

因为由函数矩阵的乘积和逆矩阵的微分法，即上章的（3-1?）式和（3-19）式，就有

$$\frac{d}{dt}\left[YX^{-1}\right]=\dot{Y}X^{-1}-YX^{-1}\dot{X}X^{-1}$$

$$=DXX^{-1}+AYX^{-1}-YX^{-1}(-BX-CY)X^{-1}$$

$$=A(YX^{-1})+(YX^{-1})B+(YX^{-1})C(YX^{-1})+D$$

可见YX^{-1}满足方程（4-37）。

2° YX^{-1}满足边界条件$P(t_0)=E$

因为X与Y是方程（4-38）的解，故有

$$X(t_0)=I,\ Y(t_0)=E$$

从而有

$$P(t_0)=Y(t_0)X^{-1}(t_0)=EI=E.$$

容易看出(4-38)式为一个齐次矩阵微分方程。据前节第一部分可知，其满足边界条件的解为：

$$W(t)=\Phi(t,\ t_0)W(t_0) \tag{4-39}$$

其中：$\Phi(t,\ t_0)$为$2n\times 2n$阶的状态转移矩阵，且满足

$$\dot{\Phi}(t,\ t_0)=G(t)\Phi(t,\ t_0) \tag{4-40}$$

及 $\quad \Phi(t_0,\ t_0)=I_{2n}$

若记

$$\Phi(t,\ t_0)=\begin{bmatrix} Q_{11}(t,\ t_0) & Q_{12}(t,t_0) \\ Q_{21}(t,\ t_0) & Q_{22}(t,\ t_0) \end{bmatrix} \tag{4-41}$$

则方程（4-38）满足边界条件的解为：

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix}=\begin{bmatrix} Q_{11}(t,\ t_0) & Q_{12}(t,t_0) \\ Q_{21}(t,\ t_0) & Q_{22}(t,t_0) \end{bmatrix}\begin{bmatrix} I \\ E \end{bmatrix}$$

$$=\begin{bmatrix} Q_{11}(t,t_0)+Q_{12}(t,t_0)E \\ Q_{21}(t,t_0)+Q_{22}(t,t_0)E \end{bmatrix} \tag{4-42}$$

于是Riccati方程(4-37)满足边界条件的解为：

$$P(t)=[Q_{21}(t,t_0)+Q_{22}(t,t_0)E][Q_{11}(t,\ t_0)+Q_{12}(t,t_0)E]^{-1} \tag{4-43}$$

可见，求解Riccati方程，关键是求解(4-38)的状态转移矩阵$\Phi(t,t_0)$，由前面可知，当矩阵G为常数矩阵而$\Phi(t,t_0)$成为指数矩阵函数e^{Gt}时，就可求得方程(4-37)的解析解。

〔**例**〕 求解下列矩阵Riccati方程

$$\dot{P}(t)=\begin{bmatrix} 0 & -1 \\ 0 & \sqrt{2} \end{bmatrix}P(t)+P(t)\begin{bmatrix} 0 & 0 \\ -1 & \sqrt{2} \end{bmatrix}$$

$$+P(t)\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}P(t)+\begin{pmatrix}0 & 0\\0 & -1\end{pmatrix}P(0)=0$$

解 由(4-40)式，有

$$\dot{\Phi}(t,0)=G\Phi(t,0) \tag{4-44}$$

$$G=\begin{pmatrix}0 & 0 & -1 & 0\\1 & -\sqrt{2} & 0 & 0\\0 & 0 & 0 & -1\\0 & -1 & 0 & \sqrt{2}\end{pmatrix} \tag{4-45}$$

据前节第一部分可知

$$\Phi(t,\ 0)=e^{Gt} \tag{4-46}$$

对于指数矩阵函数e^{Gt}的计算，需把G化为Jordan 标 准形(或者用矩阵函数的定义；或者用矩阵函数的基本公式。)

通过计算，知G的特征值$\lambda_{1,2}=1$(二重)；$\lambda_{3,4}=-1$(二重)不变因子：$d_1=1$；$d_2=1$；$d_3=1$；$d_4=(\lambda-1)^2(\lambda+1)^2$则$G$的Jordan标准形为：

$$J=\begin{pmatrix}1 & 1 & & \\ & 1 & & \\ & & -1 & 1\\ & & & -1\end{pmatrix} \tag{4-47}$$

对应于$\lambda=1$，-1的特征向量

$$\eta_1=\begin{pmatrix}1\\ \sqrt{2}-1\\ -1\\ 1\end{pmatrix};\qquad \eta_3=\begin{pmatrix}1\\ \sqrt{2}+1\\ 1\\ 1\end{pmatrix}$$

由公式：

$$(I-A)\eta_2=\eta_1$$

$$(-I-A)\eta_4=-\eta_3$$

解出

$$\eta_2=\begin{bmatrix}-1\\ \sqrt{2}-2\\ 0\\ 1\end{bmatrix};\qquad \eta_4=\begin{bmatrix}1\\ -\sqrt{2}-2\\ 0\\ -1\end{bmatrix}$$

因此得到将A化为Jordan标准形的相似变换矩阵：

$$P=[\eta_1\eta_2\eta_3\eta_4]=\begin{bmatrix}1 & -1 & 1 & 1\\ \sqrt{2}-1 & \sqrt{2}-2 & \sqrt{2}+1 & -\sqrt{2}-2\\ -1 & 0 & 1 & 0\\ 1 & 1 & 1 & -1\end{bmatrix} \tag{4-48}$$

$$P^{-1}=\frac{1}{4}\begin{bmatrix}1 & 0 & -2 & 1\\ -1 & -1 & 1 & \sqrt{2}+1\\ 1 & 0 & 2 & 1\\ 1 & -1 & 1 & \sqrt{2}-1\end{bmatrix} \tag{4-49}$$

从上章(3-54)式知：

$$e^{Jt}=\begin{bmatrix}e^{t} & te^{t} & & \\ & e^{t} & & \\ & & e^{-t} & te^{-t}\\ & & & e^{-t}\end{bmatrix} \tag{4-50}$$

再从上章（3-53）式，得

$$e^{Gt}=Pe^{Jt}P^{-1}$$

从而由(4-46)式，可得状态转移矩阵

$$\Phi(t,\ 0)=e^{Gt}=Pe^{Jt}P^{-1}$$

$$
=\frac{1}{4}\begin{bmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & \varphi_{14} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} & \varphi_{24} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} & \varphi_{34} \\ \varphi_{41} & \varphi_{42} & \varphi_{43} & \varphi_{44} \end{bmatrix}
$$

其中：

$$\varphi_{11}=2e^{t}-te^{t}+2e^{-t}+te^{-t}$$

$$\varphi_{12}=e^{t}-te^{t}-e^{-t}-te^{-t}$$

$$\varphi_{13}=-3e^{t}+te^{t}+3e^{-t}+te^{-t}$$

$$\varphi_{14}=-\sqrt{2}e^{t}+(1+\sqrt{2})te^{t}+\sqrt{2}e^{-t}+(\sqrt{2}-1)te^{-t}$$

$$\varphi_{21}=e^{t}-(\sqrt{2}-1)te^{t}-e^{-t}+(\sqrt{2}+1)te^{-t}$$

$$\varphi_{22}=(2-\sqrt{2})e^{t}-(\sqrt{2}-1)te^{t}+(2+\sqrt{2})e^{-t}-(\sqrt{2}+1)te^{-t}$$

$$\varphi_{23}=-\sqrt{2}e^{t}+(\sqrt{2}-1)te^{t}+\sqrt{2}e^{-t}+(\sqrt{2}+1)te^{-t}$$

$$\varphi_{24}=-e^{t}+te^{t}+e^{-t}+te^{-t}$$

$$\varphi_{31}=-e^{t}+te^{t}+e^{-t}+te^{-t}$$

$$\varphi_{32}=te^{t}-te^{-t}$$

$$\varphi_{33}=2e^{t}-te^{t}+2e^{-t}+te^{-t}$$

$$\varphi_{34}=-e^{t}-(1+\sqrt{2})te^{t}+e^{-t}+(\sqrt{2}-1)te^{-t}$$

$$\varphi_{41}=-te^{t}+te^{-t}$$

$$\varphi_{42}=-e^{t}-te^{t}+e^{-t}-te^{-t}$$

$$\varphi_{43}=-e^{t}+te^{t}+e^{-t}+te^{-t}$$

$$\varphi_{44}=(2+\sqrt{2})e^{t}+(\sqrt{2}+1)te^{t}+(2-\sqrt{2})e^{-t}+(\sqrt{2}-1)te^{-t}$$

再由公式(4-43)，并注意到$P(0)=E=0$，就得到所给Riccati方程的解：

$$
P(t)=\frac{1}{4}\begin{bmatrix} \varphi_{31} & \varphi_{32} \\ \varphi_{41} & \varphi_{42} \end{bmatrix}\left(\frac{1}{4}\begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}\right)^{-1}
$$

$$=\begin{pmatrix}\varphi_{31} & \varphi_{32}\\ \varphi_{41} & \varphi_{42}\end{pmatrix}\begin{pmatrix}\varphi_{11} & \varphi_{12}\\ \varphi_{21} & \varphi_{22}\end{pmatrix}^{-1}$$

习 题 四

1． 求下列微分方程组的通解。

(1) $\begin{cases}\dfrac{dx_1}{dt}=x_2+x_3\\ \dfrac{dx_2}{dt}=x_1+x_2-x_3\\ \dfrac{dx_3}{dt}=x_2+x_3\end{cases}$　　(2) $\begin{cases}\dfrac{dx_1}{dt}=x_2\\ \dfrac{dx_2}{dt}=-x_1\end{cases}$

(3) $\begin{cases}\dfrac{ds_1}{dt}=-s_1+s_2\\ \dfrac{ds_2}{dt}=-4s_1+3s_2\\ \dfrac{ds_3}{dt}=s_1+2s_3\end{cases}$

2． 用状态转移矩阵求解下列矩阵微分方程。

(1) 求 $\dfrac{dX}{dt}=AX$

满足初始条件 $X(0)=X_0$ 的特解。

其中： $X=\begin{pmatrix}x_1(t)\\ x_2(t)\end{pmatrix}$， $A=\begin{pmatrix}-5 & 1\\ 1 & -3\end{pmatrix}$， $X_0=\begin{pmatrix}1\\ 0\end{pmatrix}$

(2) 求 $\dfrac{dX}{dt}=AX+Bu(t)$

满足条件　　$X(0)=X_0$

的解。

其中：

(a) $X=\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$, $A=\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$, $B=\begin{pmatrix} -1 \\ 4 \end{pmatrix}$; $u(t)=e^{-t}$;

$$X_0=\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

(b) $X=\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}$; $A=\begin{pmatrix} -6 & 1 & 0 \\ -11 & 0 & 1 \\ -6 & 0 & 0 \end{pmatrix}$; $B=\begin{pmatrix} 2 \\ 6 \\ 2 \end{pmatrix}$

$$u(t)=[1]; \quad X_0=\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

(3) 求 $\dfrac{dx}{dt}A(t)X$

满足初始条件 $X(0)=\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

的解。

其中： $A(t)=\begin{pmatrix} t & 1 \\ 2 & t \end{pmatrix}$; $X=\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$

3. 求下列矩阵微分方程

(1) $\dfrac{d}{dt}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=\begin{pmatrix} t & 1 \\ 2 & t \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}+\begin{pmatrix} 0 \\ t \end{pmatrix}$

(2) $\dfrac{d}{dt}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}=\begin{pmatrix} 0 & 1 \\ -e^t & t^2 \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}+\begin{pmatrix} 0 \\ t^3 \end{pmatrix}$

满足条件

$$X(0)=\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

的解。

第五章　广义逆矩阵及其应用

在线性代数中讨论过一个矩阵A的逆矩阵问题。当A是满秩（非奇异）时，则A^{-1}存在，并有关系式

$$A^{-1}A=AA^{-1}=I \tag{5-1}$$

成立。此时线性方程组

$$AX=B \tag{5-2}$$

的解　$X=A^{-1}B$是唯一的。

然而，非奇异矩阵只是矩阵的一种特殊情况。实际上，我们经常遇到的是奇异矩阵或长矩阵。这些最一般的矩阵显然不存在满足（5-1）式的那种逆矩阵A^{-1}，而方程组(5-2）的解存在的充分必要条件是

$$\text{rank}[A,\ B]=\text{rank}A$$

这时，其解是否也可能用某个矩阵G表示为

$$X=GB \tag{5-3}$$

呢？这个问题的回答是肯定的，我们可以把逆矩阵的概念加以推广，使对任意$m\times n$阶矩阵A，一般地$m\neq n$，且A有任意秩都可以建立广义逆矩阵的概念。当方程组（5-2）为相容时，其解可以表示成（5-3）式形式；而当方程组（5-2）为不相容时，仍可以用（5-3)式的形式表示该矛盾方程组在一定意义下的最优近似解。因此，广义逆矩阵的概念不仅与线性方程组的求解问题有关，也在最优控制问题中有用。

广义逆矩阵的理论已成为数理统计、最优化理论、现代控制理论和网络理论等学科的重要工具，是矩阵理论在最近三十多年中的新成就之一。早在1920年，Moore首先提出奇

异矩阵的逆矩阵问题，从而建立了广义逆矩阵的概念。后来在1935年他又对广义逆矩阵的概念进行了讨论。1955年Penrose也独立提出了一个广义逆矩阵的概念。实际上，这两种广义逆矩阵的定义是等价的，因此称为 Moore-Penrose 广义逆，记作A^+。同年Rao提出了一个更一般的广义逆矩阵的概念，称为g逆，记作A^-。广义逆矩阵的理论在此基础上得到充分发展，目前已广泛用于许多学科中。

本章着重讨论A^-与A^+及其在方程组求解中的应用。

第一节 广义逆矩阵A^-的概念

一、满秩长矩阵的右逆和左逆

在讨论广义逆矩阵A^-之前，首先谈谈满秩长矩阵的右逆，左逆的概念。

定义1 设矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{5-4}$$

如果，当$m\leqslant n$时，有$\text{rank}A=m$；或者当$m\geqslant n$时，有$\text{rank}A=n$，则称这两种长矩阵为满秩长矩阵，前者称为**行满秩**，后者称为**列满秩**。

可以证明满秩长矩阵具有如下性质。

性质1 $\text{rank}(AA^T)=m \quad (A=[a_{ij}]_{m\times n};\ m\leqslant n)$ (5-5)

性质2 $\text{rank}(A^TA)=n \quad (A=[a_{ij}]_{m\times n};\ m\geqslant n)$ (5-6)

定义2　设A是行满秩的$m\times n$阶实长矩阵（$m\leqslant n$），若存在一个$n\times m$阶矩阵G，当G右乘A后得到一个$m\times m$阶单位矩阵I，即

$$AG=I \tag{5-7}$$

则G称为A的右逆，并记作A_R^{-1}。即有

$$AA_R^{-1}=I \tag{5-8}$$

又由（5-5）式知，乘积AA^T是普通的满秩方阵，故有：

$$(AA^T)^{-1}(AA^T)=(AA^T)(AA^T)^{-1}=I \tag{5-9}$$

由右逆定义，得

$$A_R^{-1}=A^T(AA^T)^{-1} \tag{5-10}$$

定义3　设A是列满秩的$m\times n$阶实长矩阵($m\geqslant n$)，若存在一个$n\times m$阶矩阵G，当G左乘A后得到一个$n\times n$阶单位阵，即

$$GA=I \tag{5-11}$$

则G称为A的**左逆**，并记作A_L^{-1}。即有

$$A_L^{-1}A=I \tag{5-12}$$

又由（5-6）式知，乘积A^TA是普通的满秩方阵，故有：

$$(A^TA)(A^TA)^{-1}=(A^TA)^{-1}(A^TA)=I \tag{5-13}$$

由左逆定义，得

$$A_L^{-1}=(A^TA)^{-1}A^T \tag{5-14}$$

必须指出：

1°　对于行（或列）满秩的$m\times n$阶矩阵A，在$m\neq n$时A_R^{-1}和A_L^{-1}不可能同时存在。仅当$m=n$时，A_R^{-1}和A_L^{-1}才能同时存在，且等于普通的逆矩阵A^{-1}。

2°　行（或列）满秩矩阵的右逆（或左逆）并不是唯一的，(5-10)式（或(5-14)式）仅表示所有右逆（或左逆）中的一个。下面给出矩阵A的右逆（或左逆）的一般表达式。

定理1 设矩阵 $A=[a_{ij}]_{m\times n}$，且$\mathrm{rank}A=m$，则矩阵A的右逆的一般表达式是

$$G=VA^T(AVA^T)^{-1} \tag{5-15}$$

其中：V是使等式$\mathrm{rank}\ (AVA^T)=\mathrm{rank}A$成立的任意$n$阶方阵。

证明 用A左乘（5-15）式两边，得

$$AG=AVA^T(AVA^T)^{-1}$$

由于$\mathrm{rank}(AVA^T)=\mathrm{rank}A=m$，所以$(AVA^T)$是满秩方阵。因此，有

$$(AVA^T)(AVA^T)^{-1}=I$$

于是

$$G=VA^T(AVA^T)^{-1}$$

就是A的右逆的一般表达式。

由于满足$\mathrm{rank}(AVA^T)=\mathrm{rank}A=m$的矩阵$V$不是唯一的，所以右逆$G$也不是唯一的。

还应指出：若当$V=I_{n\times n}$时，（5-15）式就变成了（5-10）式。所以，（5-10）式所表示的右逆A_R^{-1}是$G=VA^T(AVA^T)^{-1}$中的一个。

同理，可以给出矩阵A的左逆的一般表达式：

$$G=(A^TUA)^{-1}A^TU \tag{5-16}$$

其中：U是满足关系式$\mathrm{rank}(A^TUA)=\mathrm{rank}A=n$的任意$m$阶方阵。

〔例1〕 设矩阵

$$A=\begin{bmatrix}1 & 2 & -1\\ 0 & -1 & 2\end{bmatrix}$$

求A的右逆。

解 显然$\mathrm{rank}A=2$，即A是行满秩矩阵。由(5-10)式，

可得

$$A_R^{-1}=A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}=\begin{bmatrix}1&0\\2&-1\\-1&2\end{bmatrix}\left(\begin{bmatrix}1&2&-1\\0&-1&2\end{bmatrix}\begin{bmatrix}1&0\\2&-1\\-1&2\end{bmatrix}\right)^{-1}$$

$$=\begin{bmatrix}1&0\\2&-1\\-1&2\end{bmatrix}\begin{bmatrix}6&-4\\-4&5\end{bmatrix}^{-1}=\frac{1}{14}\begin{bmatrix}5&4\\6&2\\3&8\end{bmatrix}$$

若选择

$$V=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}$$

V显然满足$\operatorname{rank}(AVA^{\mathrm{T}})=\operatorname{rank}A=2$，故由(5-15)式，可求出$A$的另一个右逆：

$$G=VA^{\mathrm{T}}(AVA^{\mathrm{T}})^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0\\2&-1\\-1&2\end{bmatrix}$$

$$\times\left(\begin{bmatrix}1&2&-1\\0&-1&2\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0\\2&-1\\-1&2\end{bmatrix}\right)^{-1}$$

$$=\begin{bmatrix}1&0\\2&-1\\0&0\end{bmatrix}\begin{bmatrix}5&-2\\-2&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0\\2&-1\\0&0\end{bmatrix}\begin{bmatrix}1&2\\2&5\end{bmatrix}$$

$$=\begin{bmatrix}1&2\\0&-1\\0&0\end{bmatrix}$$

〔例2〕 设矩阵

$$A=\begin{pmatrix}1 & 2\\ 2 & 1\\ 1 & 1\end{pmatrix}$$

求A的左逆。

解 显然rank$A=2$，即A是一个列满秩矩阵，由(5-14)式，可求得

$$A_L^{-1}=(A^TA)^{-1}A^T=\left(\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & 1\\ 1 & 1\end{pmatrix}\right)^{-1}\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}$$

$$=\begin{pmatrix}6 & 5\\ 5 & 6\end{pmatrix}^{-1}\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}=\frac{1}{11}\begin{pmatrix}-4 & 7 & 1\\ 7 & -4 & 1\end{pmatrix}$$

若选择

$$U=\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}$$

显然rank$(A^TUA)=$rank$A=2$，再由(5-16)式，可求得A的另一个左逆：

$$G=(A^TUA)^{-1}A^TU=\left(\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & 2\\ 2 & 1\\ 1 & 1\end{pmatrix}\right)^{-1}$$

$$\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$$=\begin{pmatrix}-3 & 0\\ 0 & 3\end{pmatrix}^{-1}\begin{pmatrix}1 & -2 & 0\\ 2 & -1 & 0\end{pmatrix}=\frac{1}{3}\begin{pmatrix}-1 & 2 & 0\\ 2 & -1 & 0\end{pmatrix}$$

对于满秩的长矩阵A，当其右逆为$A_R^{-1}=A^T(AA^T)^{-1}$(或左逆$A_L^{-1}=(A^TA)^{-1}A^T$)时，显然还满足下列四个等式：

(1) $AA_R^{-1}A=A$ (或$AA_L^{-1}=A$) (5-17)

(2) $A_R^{-1}AA_R^{-1}=A_R^{-1}$(或$A_L^{-1}AA_L^{-1}=A_L^{-1}$) (5-18)

(3) $(AA_R^{-1})^T=AA_R^{-1}$(或$(AA_L^{-1})^T=AA_L^{-1}$ (5-19)

(4) $(A_R^{-1}A)^T=A_R^{-1}A$(或$(A_L^{-1}A)^T=A_L^{-1}A$ (5-20)

但是右逆（或左逆）的一般表达式$G=VA^T(AVA^T)^{-1}$(或$G=(A^TUA)^{-1}A^TU$)只能满足(5-17)式和(5-18)式。

二、广义逆矩阵A^-的定义及其一般表达式

前面我们已对满秩长矩阵A的右逆（或左逆）进行了讨论。下面还要讨论非满秩长矩阵A的逆矩阵问题。

定义4 设A为$m\times n$阶实矩阵，且$\mathrm{rank}A=r\leqslant\min(m,n)$，若存在$n\times m$阶矩阵$A^-$，满足矩阵方程

$$AA^-A=A \tag{5-21}$$

则称A^-为矩阵A的广义逆矩阵，也叫做**g逆**。

下面将证明满足（5-21）式的g逆不但存在，而且一般说来不是唯一的。

若当$\mathrm{rank}A=r=m$（或n)时，即矩阵A是行（或列）满秩的，显然A_R^{-1}（或A_L^{-1}）就是A的一个g逆。因为A_R^{-1}（或A_L^{-}）显然满足(5-21)式。

若当$\mathrm{rank}A=r<\min(m,n)$时，这类非满秩长矩阵总存在一个**满秩分解**，即是总存在$m\times r$阶的列满秩阵C和$r\times n$阶的行满秩阵D，使得

$$A=CD$$

因此，显然存在C的左逆C_L^{-1}和D的右逆D_R^{-1}，而且

$$A^-=D_R^{-1}C_L^{-1} \tag{5-22}$$

这就是A的一个g逆。

(5-22)式可用定义4验证，即

$$AA^{-}A=AD_R^{-1}C_L^{-1}A=CDD_R^{-1}C_L^{-1}CD=CD=A$$

这样，(5-22)式就具体地给出了构造A的g逆的方法。下面我们将进一步给出g逆的一般表达式，并阐明g逆的不唯一性。

定理2 设矩阵$A=[a_{ij}]_{m\times n}$，A^{-}是A的一个广义逆，且$\text{rank}A\leqslant\min(m,n)$，则$A$的$g$逆的一般表达式为：

$$G=A^{-}+U-A^{-}AUAA^{-} \tag{5-23}$$

或

$$G=A^{-}+V(I-AA^{-})+(I-A^{-}A)W \tag{5-24}$$

其中：A^{-}是A的任一个g逆，而U，V，W是具有适当阶数的任意矩阵。

证明 由(5-21)式，可得

$$\begin{aligned}AGA&=A(A^{-}+U-A^{-}AUAA^{-})A=AA^{-}A+AUA\\&\quad-AA^{-}AUAA^{-}A=A+AUA-AUA=A\end{aligned}$$

或

$$\begin{aligned}AGA&=A(A^{-}+V(I-AA^{-})+(I-A^{-}A)W)A\\&=AA^{-}A+AVA-AVAA^{-}A+AWA-AA^{-}AWA\\&=A+AVA-AVA+AWA-AWA=A\end{aligned}$$

从而，就证明了(5-23)式和(5-24)式都是A的g逆。由于U，V，W可以任意选择，故g逆不是唯一的。

下面再证明(5-23)式，(5-24)式都是A的g逆的一般表达式。

设G是A的任何一个g逆，有

$$O=A(G-A^{-})A\Longrightarrow G=A^{-}+(G-A^{-})-A^{-}A(G-A^{-})AA^{-}$$

令 $U=G-A^{-}$，即得

$$G=A^{-}+U-A^{-}AUAA^{-}$$

恒成立。

又(5-23)式和(5-24)式显然是等价的。

事实上，若令 $U=V(I-AA^-)+(I-A^-A)W$，于是有

$$AUA=AVA-AVAA^-A+AWA-AA^-AWA$$
$$=AVA-AVA+AWA-AWA=0$$

所以

$$A^-+V(I-AA^-)+(I-A^-A)W$$
$$=A^-+U=A^-+U-A^-AUAA^-$$

这就证明了可由(5-24)式，得出(5-23)式。

反之，由于

$$A^-+U-A^-AUAA^-$$
$$=A^-+U(I-AA^-)+(I-A^-A)UAA^-$$

取$V=U$，$W=UAA^-$，则由(5-23)式得出(5-24)式。从而就证明(5-23)式与(5-24)式的等价性。由此也证明了(5-24)式也是A的g逆的一般表达式。

如果A是满秩方阵，则有$A^-=A^{-1}$，此时(5-23)式和(5-24)式都只能得出唯一的普通逆。由此可见，广义逆矩阵确是普通逆矩阵概念的推广，而普通逆矩阵是广义逆矩阵的一个特殊情况。

定理3 设矩阵

$$A=[a_{ij}]_{m\times n};\quad B=[b_{ij}]_{r\times n}$$

A^-是矩阵A的一个广义逆矩阵。等式：

$$BA^-A=B \tag{5-25}$$

成立的充要条件是存在一个矩阵D，使得满足

$$B=DA \tag{5-26}$$

证明 充分性是显然的。必要性如下证明：

设 $BA^-A=B$， 取 $D=BA^-$且右乘A

则 $DA=BA^-A$

所以 $DA=B$

这就证明了使（5-25）式成立的矩阵D是存在的。

进一步阐明：

1° 若$B=[b_{ij}]_{m\times r}$，则（5-25）式就变为

$$AA^{-}B=B$$

相应地（5-26）式就变为：

$$B=AD$$

2° 设

$$B=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix} \text{其中：} b_i=[b_{i1}\ b_{i2}\ \cdots\ b_{in}] \quad (i=1,\ 2,\ \cdots,\ r)$$

$$A=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \text{其中：} a_i=[a_{i1}\ a_{i2}\ \cdots\ a_{in}] \quad (i=1,\ 2,\ \cdots,\ m)$$

$$D=\begin{pmatrix} d_{11} & d_{12} & \cdots & d_{1m} \\ d_{21} & d_{22} & \cdots & d_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ d_{r1} & d_{r2} & \cdots & d_{rm} \end{pmatrix}$$

由（5-26）式，有

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{pmatrix}=\begin{pmatrix} d_{11} & d_{12} & \cdots d_{1m} \\ d_{21} & d_{22} & \cdots d_{2m} \\ \cdots & \cdots & \cdots \\ d_{r1} & d_{r2} & \cdots d_{rm} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

所以 $$b_i=\sum_{j=1}^{n} d_{ij}a_j \qquad (i=1,\ 2,\ \cdots,\ r)$$

说明B的所有行向量都是A的行向量的线性组合。类似地，$B=AD$，表示B的所有列向量都是A的列向量的线性组合。因此，定理3就等价于下列定理。

定理4 等式$BA^-A=B$成立的充要条件是：B的所有行向量都是A的行向量的线性组合。（或者说：等式$AA^-B=B$成立的充要条件是：B的所有列向量都是A的列向量的线性组合）。

定理3和定理4说明：在一定条件下，AA^-与A^-A"就象"一个普通的单位矩阵。即是说：A^-"就象"是A的普通逆矩阵。

三、广义逆矩阵A^-的性质

性质1 $(A^-)^T=(A^T)^-$ (5-27)

即A的g逆的转置与A的转置的g逆相等。

证明 由定义4 $AA^-A=A$

于是，有 $(AA^-A)^T=A^T$，即$A^T(A^-)^TA^T=A^T$

故 $(A^T)^-=(A^-)^T$

性质2 $A(A^TA)^-A^TA=A$ (5-28a)

或 $A^TA(A^TA)^-A^T=A^T$ (5-28b)

即 $(A^TA)^-A^T$是A的一个g逆，或$A(A^TA)^-$是A^T的一个g逆。

证明 因为：

$$[A(A^TA)^-A^TA-A]^T[A(A^TA)^-A^TA-A]$$
$$=[A^TA[(A^TA)^-]^TA^T-A^T][A(A^TA)^-A^TA-A]$$
$$=[A^TA(A^TA)^--I][A^TA(A^TA)^-A^TA-A^TA]$$
$$=[A^TA(A^TA)^--I][A^TA-A^TA]=0$$

由于对任意矩阵C，如果$C^TC=0$，则必有$C=0$。故有

$$A(A^TA)^-A^TA - A = 0$$

从而得到（5-28a）式

$$A(A^TA)^-A^TA = A$$

对（5-28a）式两端转置，得

$$A^TA(A^TA)^-A^T = A^T$$

性质3 $AGA = A$的充要条件是：

$$A^TAGA = A^TA$$

证明 必要性：设$AGA = A$，显然可得出

$$A^TAGA = A^TA$$

充分性：设$A^TAGA = A^TA$，于是有

$$A^TAGA - A^TA = 0$$

上式两边左乘 $(GA)^T - I$，得

$$[(GA)^T - I][A^TAGA - A^TA] = 0$$

$$[(GA)^T - I]A^T[AGA - A] = 0$$

$$[(GA)^TA^T - A^T][AGA - A] = 0$$

$$[AGA - A]^T[AGA - A] = 0$$

有 $AGA - A = 0$

故 $AGA = A$

性质4 $\operatorname{rank}(A^-) \geqslant \operatorname{rank}A$ （5-29）

即是A^-的秩不小于A的秩。

证明 根据线性代数知识，设矩阵

$$A = CD$$

则有 $\operatorname{rank}A \leqslant \operatorname{rank}C$和$\operatorname{rank}A \leqslant \operatorname{rank}D$

由于 $AA^-A = A$

于是有 $\operatorname{rank}(AA^-) \geqslant \operatorname{rank}A$，又因有

$$\operatorname{rank}(A^-) \geqslant \operatorname{rank}(AA^-),$$

所以 $\operatorname{rank}(A^-) \geqslant \operatorname{rank}(AA^-) \geqslant \operatorname{rank}A$

故　　$\text{rank}(A^-)\geqslant \text{rank}A$

应该注意：对普通的满秩方阵A，有$(A^{-1})^{-1}=A$，但对于g逆一般不存在这种关系式，就是说对奇异方阵或长矩阵，A本身不是g逆的g逆。即$(A^-)^-\neq A$。但是，对于g逆的一个子集可能具有这种反射性质。下面给出"反射g逆"的定义。

定义5　设G是$n\times m$阶矩阵，如果满足

$$AGA=A\text{和}GAG=G$$

则称G为A的一个**反射g逆**，记作

$$G=A_r^-$$

显然，它是g逆的一个子集合。

由于A的反射g逆是对g逆用条件$GAG=G$限制而得出的，故满足反射性质$(A^-)^-=A$

行满秩阵的右逆和列满秩阵的左逆都是A的一个反射g逆。

如果已知A的任一个g逆A^-，则$G=A^-AA^-$是A的一个反射g逆。如果G_1，G_2分别是A的两个不同的g逆，则G_1AG_2是A的一个反射g逆。

四、广义逆矩阵A^-的计算方法

1°　满秩分解法

为了解决广义逆矩阵的计算问题，先介绍矩阵的满秩分解，它是矩阵理论的基础知识。

定理5　设矩阵$A=[a_{ij}]_{m\times n}$，且$\text{rank}A=r<\min(m,n)$，则经过有限次初等行变换可把A化成

$$\tilde{A}_r=\begin{pmatrix} 0 & \cdots & 0 & \overset{(k_1)}{1} & * & \cdots & * & \overset{(k_2)}{0} & * & \cdots & * & \overset{(k_r)}{0} & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}\begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} r\text{行} \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} m-r\text{行} \end{matrix}$$

其中：$1\leqslant k_1<k_2<\cdots<k_r\leqslant n$，$*$号的元素不一定是零。

$\tilde{A}_r$中第k_i个列向量为：

$$e_{k_i}=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\text{第}i\text{个分量}\quad (i=1,\ 2,\ \cdots,\ r)$$

证明 设$\operatorname{rank}A=r\geqslant 1$

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}=[a_1\ \ a_2\ \ \cdots\ \ a_n]$$

如果A的第一个非零列向量为：

$$a_{k_1}=\begin{pmatrix} a_{1k_1} \\ a_{2k_1} \\ \vdots \\ a_{ik_1} \\ \vdots \\ a_{mk_1} \end{pmatrix}\neq 0$$

不妨设$a_{ik_1}\neq 0$，将第一行和第i行互换后，再把第一行乘以$\frac{1}{a_{ik_1}}$，得

$$\begin{pmatrix} & & \overset{(k_1)}{} & & & \\ 0 & \cdots 0 & 1 & * & \cdots & * \\ 0 & \cdots 0 & a_{2k} & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots 0 & a_{1k_1} & * & \cdots & * \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots 0 & a_{mk_1} & * & \cdots & * \end{pmatrix} \text{第}i\text{行}$$

用$-a_{jk_1}$乘第一行加到第j行（$j=1, 2, \cdots, m$；$j\neq i$）又用$-a_{1k_1}$乘第一行加到第i行，得

$$\begin{pmatrix} & & \overset{(k_1)}{} & & \\ 0 & \cdots 0 & 1 & * \cdots * \\ 0 & \cdots 0 & 0 & * \cdots * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots 0 & 0 & * \cdots * \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots 0 & 0 & * \cdots * \end{pmatrix} \quad \text{令} A_1 = \begin{pmatrix} * \cdots * \\ \cdots\cdots \\ * \cdots * \\ \cdots\cdots \\ * \cdots * \end{pmatrix}_{(m-1)\times(n-k_1)}$$

若$A_1=0$，则定理得证。若$A_1\neq 0$，可再用此法化为

$$\begin{pmatrix} 0 \cdots 0 & 1 & * \cdots * & 0 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 1 & * \cdots * \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & * \cdots * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & * \cdots * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 \cdots 0 & 0 & 0 \cdots 0 & 0 & * \cdots * \end{pmatrix}$$

如此继续使用此法，经过有限次初等行变换后就能把 A 化为 $\widetilde{A}$ 。

定理6 设 $A=[a_{ij}]_{m\times n}$，且 $\operatorname{rank}A=r<\min(m,n)$，则可将 A 作满秩分解（或称 A 的最大秩分解）：

$$A=CD$$

其中：C 是 $m\times r$ 阶矩阵，D 是 $r\times n$ 阶矩阵，且

$$\operatorname{rank}C=\operatorname{rank}D=r$$

证明 由定理5，A 通过有限次初等行变换可化成 $\widetilde{A}_r$，则 $\widetilde{A}_r$ 中的 k_1，$k_2,\cdots,k_r$ 个列向量 $\widetilde{a}_{k_1}$，$\widetilde{a}_{k_2},\cdots,\widetilde{a}_{k_r}$ 线性无关。因此，有

$$\widetilde{a}_j=\begin{pmatrix} l_{1j}\\ l_{2j}\\ \vdots\\ l_{pj}\\ 0\\ \vdots\\ 0\end{pmatrix}=l_{1j}\widetilde{a}_{k_1}+l_{2j}\widetilde{a}_{k_2}+\cdots+l_{pj}\widetilde{a}_{k_p}$$

其中：$1\leqslant p\leqslant r$；$j=1,\ 2,\ \cdots,\ n$

于是 $a_j=l_{1j}a_{k_1}+l_{2j}a_{k_2}+\cdots+l_{pj}a_{k_p}$

取 $C=[a_{k_1},\ a_{k_2},\ \cdots,\ a_{k_r}]_{m\times r}$

则 $\operatorname{rank}C=r$，把 $\widetilde{A}_r$ 中的前 r 行取为 D，

即

$$D=\begin{pmatrix} 0\cdots 0 & 1 & *\cdots * & 0 & *\cdots * & 0 & *\cdots *\\ 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * & 0 & *\cdots *\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0\cdots 0 & 0 & 0\cdots 0 & 0 & 0\cdots 0 & 1 & *\cdots * \end{pmatrix}_{r\times n}$$

则有 $\operatorname{rank}D=r$，且 D 中的第 j 个列向量

$$b_j=\begin{bmatrix} l_{1j} \\ l_{2j} \\ \vdots \\ l_{pj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}\begin{matrix} \left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\} p\text{行} \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} r-p\text{行} \end{matrix} \qquad (1\leqslant p\leqslant r;\ j=1,\ 2,\ \cdots,\ n)$$

于是 $$Cb_j=[a_{k_1}\ a_{k_2}\ \cdots\ a_{k_r}]\begin{bmatrix} l_{1j} \\ l_{2j} \\ \vdots \\ l_{pj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$=l_{1j}a_{k_1}+l_{2j}a_{k_2}+\cdots+l_{pj}a_{k_p}=a_j$$

故 $$CD=C[b_1\ b_2\ \cdots\ b_n]=[Cb_1\ Cb_2\ \cdots Cb_n]$$

$$=[a_1\ a_2\ \cdots\ a_n]=A$$

举例说明如何求矩阵C和D

〔例3〕 设$A=\begin{bmatrix} 2 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 2 & -1 & 1 & 2 \end{bmatrix}$

求矩阵C和D，使$A=CD$，且$\text{rank}A=\text{rank}C=\text{rank}D$

解 对A用初等行变换

$$A=\begin{bmatrix} 2 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 2 & -1 & 1 & 2 \end{bmatrix}\simeq\begin{bmatrix} 2 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 1 & 4 \end{bmatrix}\simeq\begin{bmatrix} 2 & 0 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\simeq\begin{bmatrix}1&0&\frac{1}{2}&2\\0&1&0&2\\0&0&0&0\end{bmatrix}=\tilde{A}_r$$

则 $\text{rank}A=2$，A的一，二列线性无关

于是取 $C=\begin{bmatrix}2&0\\0&1\\2&-1\end{bmatrix}$ 有 $\text{rank}C=2$

$D=\begin{bmatrix}1&0&\frac{1}{2}&2\\0&1&0&2\end{bmatrix}$ 有 $\text{rank}D=2$

而且 $A=CD$

上面的两个定理对初等列变换也成立。因此也可用初等列变换求A的满秩分解。

〔**例4**〕 设

$$A=\begin{bmatrix}1&0&3\\2&3&0\\1&1&1\end{bmatrix}$$

求矩阵C和D，使$A=CD$，且$\text{rank}A=\text{rank}C=\text{rank}D$

解 对A用初等列变换

$$A=\begin{pmatrix}1&0&3\\2&3&0\\1&1&1\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\2&3&-6\\1&1&-2\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\2&3&0\\1&1&0\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\-1&3&0\\0&1&0\end{pmatrix}$$

$$\simeq\begin{pmatrix}1&0&0\\-1&1&0\\0&\frac{1}{3}&0\end{pmatrix}\simeq\begin{pmatrix}1&0&0\\0&1&0\\\frac{1}{3}&\frac{1}{3}&0\end{pmatrix}=\tilde{A}_r$$

则 $\text{rank}A=2$，$\tilde{A}_r$ 中的一，二列线性无关

于是取 $C=\begin{pmatrix}1&0\\0&1\\\frac{1}{3}&\frac{1}{3}\end{pmatrix}$ 有 $\text{rank}C=2$

$$D=\begin{bmatrix} 1 & 0 & 3 \\ 2 & 3 & 0 \end{bmatrix} \quad 有\mathrm{rank}D=2$$

而且 $A=CD$

注意：A的满秩因子矩阵C和D不是唯一的。还可以混合地利用初等行和列变换来解决满秩分解的问题。

设矩阵 $A=[a_{ij}]_{m\times n}$，且$\mathrm{rank}A=r<\min(m,n)$利用初等行和列变换，总可使得

$$\widetilde{A}=PAQ \tag{5-30}$$

其中：P，Q称为**初等变换阵**（起初等变换作用的矩阵）都是满秩方阵，而

$$\widetilde{A}=\begin{bmatrix} (\widetilde{A}_1)_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix}$$

由（5-30）式，可得

$$A=P^{-1}\widetilde{A}Q^{-1}=P^{-1}\begin{bmatrix} (\widetilde{A}_1)_{r\times r} & \\ & 0_{(m-r)\times(n-r)} \end{bmatrix}Q^{-1} \tag{5-31}$$

把$\widetilde{A}$分解为：

$$\widetilde{A}=\begin{bmatrix} (\widetilde{A}'_1)_{r\times r} \\ \hdashline 0_{(m-r)\times r} \end{bmatrix}[(\widetilde{A}_1'')_{r\times r} \vdots 0_{r\times(n-r)}]$$

其中：$\widetilde{A}'_1\widetilde{A}''_1=\widetilde{A}_1$，且$\widetilde{A}'_1$，$\widetilde{A}''_1$都是$r\times r$阶满秩方阵。当然分解$\widetilde{A}_1$为两个满秩阵的乘积不是唯一的。最简单的分解可令$\widetilde{A}'_1=\widetilde{A}_1$，$\widetilde{A}''=I$

这样

$$A=P^{-1}_{m\times m}\begin{bmatrix} (\widetilde{A}_1)_{r\times r} \\ \hdashline 0_{(m-r)\times r} \end{bmatrix}[I_{r\times r} \vdots 0_{r\times(n-r)}]Q^{-1}_{n\times n}$$

若取 $C_{m\times r}=P^{-1}\begin{pmatrix}\widetilde{A}_1\\ \cdots\\ 0\end{pmatrix}$； $D_{r\times n}=[I \mid 0]Q^{-1}$ (5-32)

则C，D就是A的满秩分解因子矩阵，即

$$A=CD$$

〔**例5**〕 设 $A=\begin{pmatrix}1&2&0\\0&0&2\\2&4&0\end{pmatrix}$

试用满秩分解求广义逆矩阵A^-。

解 因为rank$A=2$，对A作初等变换（行和列变换），令初等变换阵

$$P_1=\begin{pmatrix}1&0&0\\0&1&0\\-2&0&1\end{pmatrix};\qquad Q_1=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$$

则有$P_1AQ_1=\begin{pmatrix}1&0&0\\0&1&0\\-2&0&1\end{pmatrix}\begin{pmatrix}1&2&0\\0&0&2\\2&4&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}=\begin{pmatrix}1&0&2\\0&2&0\\0&0&0\end{pmatrix}$

再令 $Q_2=\begin{pmatrix}1&0&-2\\0&1&0\\0&0&1\end{pmatrix}$

则 $P_1AQ_1Q_2=\begin{pmatrix}1&0&2\\0&2&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&0&-2\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&0\\0&2&0\\0&0&0\end{pmatrix}$

所以 $\widetilde{A}=PAQ=\begin{pmatrix}1&0&0\\0&2&0\\0&0&0\end{pmatrix}$

其中： $P=P_1=\begin{pmatrix}1&0&0\\0&1&0\\-2&0&1\end{pmatrix}$；

$$Q=Q_1Q_2=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&-2\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}1&0&-2\\0&0&1\\0&1&0\end{pmatrix}$$

因此 $P^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}$； $Q^{-1}=\begin{pmatrix}1&2&0\\0&0&1\\0&1&0\end{pmatrix}$

于是 $A=P^{-1}\widetilde{A}Q^{-1}=\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&2&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&2&0\\0&0&1\\0&1&0\end{pmatrix}$

显然 $\widetilde{A}_1=\begin{pmatrix}1&0\\0&2\end{pmatrix}$

应用公式（5-32），得到A的满秩因子

$$C=P^{-1}\begin{pmatrix}\widetilde{A}_1\\ \cdots\\ 0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}\begin{pmatrix}1&0\\0&2\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\0&2\\2&0\end{pmatrix}$$

$$D=[I\ \vdots\ 0]Q^{-1}=\begin{pmatrix}1&0&0\\0&1&0\end{pmatrix}\begin{pmatrix}1&2&0\\0&0&1\\0&1&0\end{pmatrix}=\begin{pmatrix}1&2&0\\0&0&1\end{pmatrix}$$

则 $A=CD$

利用(5-10)式和(5-14)式，得

$$D_R^{-1}=\frac{1}{5}\begin{pmatrix}1&0\\2&0\\0&5\end{pmatrix}；\quad C_L^{-1}=\frac{1}{5}\begin{pmatrix}1&0&2\\0&\frac{5}{2}&0\end{pmatrix}$$

再利用(5-22)式，得

$$A^- = D_R^{-1}C_L^{-1} = \frac{1}{25}\begin{pmatrix}1 & 0\\ 2 & 0\\ 0 & 5\end{pmatrix}\begin{pmatrix}1 & 0 & 2\\ 0 & \frac{5}{2} & 0\end{pmatrix} = \frac{1}{25}\begin{pmatrix}1 & 0 & 2\\ 2 & 0 & 4\\ 0 & \frac{25}{2} & 0\end{pmatrix}$$

应当指出：用满秩分解求广义逆A^-是很麻烦的，但用此法所求的广义逆正是一个重要的广义逆——Moore-Penrose广义逆。这种特殊广义逆将在第三节中详细讨论。

2° 初等行、列变换法

设矩阵$A=[a_{ij}]_{m\times n}$，且 $\text{rank}A=r<\min(m,n)$则对A作初等行、列变换后，总可以把A写成以下分块矩阵形式：

$$\tilde{A} = \begin{pmatrix}\tilde{A}_1 & \tilde{A}_2\\ \tilde{A}_3 & \tilde{A}_4\end{pmatrix} \tag{5-33}$$

$\tilde{A}_1$是$r\times r$阶满秩方阵，$\tilde{A}_2$，$\tilde{A}_3$，$\tilde{A}_4$是满足条件

$$\tilde{A}_4 = \tilde{A}_3\tilde{A}_1^{-1}\tilde{A}_2 \tag{5-34}$$

的适当阶数的矩阵。

对A实行初等行、列变换，得

$$\tilde{A} = PAQ \tag{5-35}$$

根据广义逆定义 4 及(5-34)式，可以验证

$$\tilde{A}^- = \begin{pmatrix}\tilde{A}_1^{-1} & 0\\ 0 & 0\end{pmatrix} \tag{5-36}$$

是$\tilde{A}$的一个广义逆。

因此，由(5-35)式，有

$$A = P^{-1}\tilde{A}Q^{-1}$$

可以验证

$$A^- = (P^{-1}\tilde{A}Q^{-1})^- = Q\tilde{A}^-P$$

即 $$A^{-}=Q\begin{pmatrix}\widetilde{A}_1^{-1} & 0\\ 0 & 0\end{pmatrix}P \qquad (5\text{-}37)$$

〔例6〕 设矩阵

$$A=\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix}$$

求广义逆矩阵A^{-}。

解 由于rank$A=2$，对A作初等行、列变换，得

$$\widetilde{A}=\left(\begin{array}{cc|c}2&0&0\\0&1&1\\ \hline 1&0&0\\1&1&1\end{array}\right)$$

于是 $\widetilde{A}_1=\begin{pmatrix}2&0\\0&1\end{pmatrix}$； $\widetilde{A}^{-1}=\frac{1}{2}\begin{pmatrix}1&0\\0&2\end{pmatrix}$； $\widetilde{A}_2=\begin{pmatrix}0\\1\end{pmatrix}$；

$$\widetilde{A}_3=\begin{pmatrix}1&0\\1&1\end{pmatrix}；\quad \widetilde{A}_4=\begin{pmatrix}0\\1\end{pmatrix}$$

且满足（5-34）式。

从而 $$\widetilde{A}^{-}=\begin{pmatrix}1/2&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}$$

又 $$P=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}；\quad Q=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$$

故

$$A^{-}=Q\widetilde{A}^{-}P=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}1/2&0&0&0\\0&1&0&0\\0&0&0&0\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}0&0&0&0\\0&1&0&0\\1/2&0&0&0\end{pmatrix}$$

3° 初等行变换法

设矩阵$A=[a_{ij}]_{m\times n}$，且$\mathrm{rank}A=r<\min(m,n)$总可以作初等行变换，化$A$为如下所示分块矩阵

$$\widetilde{A}=\begin{pmatrix}C\\ —\\ B\end{pmatrix}$$

其中：C是$r\times n$阶行满秩阵，B是$(m-r)\times n$阶矩阵，且满足

$$BC_R^{-1}C=B$$

则 $$\widetilde{A}^{-}=[C_R^{-1} \vdots 0] \tag{5-38}$$

是$\widetilde{A}$的一个g逆，且具有反射性质(读者自行证明)。

对A作初等行变换等价于左乘初等变换阵P，即

$$\widetilde{A}=PA$$

于是 $A=P^{-1}\widetilde{A}$

可以验证

$$A^{-}=(P^{-1}\widetilde{A})^{-}=\widetilde{A}^{-}P \tag{5-39}$$

显然，A^{-}是A的一个g逆，且具有反射性质。

〔例7〕 设矩阵

$$A=\begin{pmatrix}0&1&0\\0&2&0\\1&0&0\end{pmatrix}$$

求广义逆矩阵A^-。

解 因为rank$A=2$，对A作初等行变换，则有

$$\widetilde{A}=\begin{pmatrix}1&0&0\\0&1&0\\ \hline 0&2&0\end{pmatrix},\quad P=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$

$$C=\begin{pmatrix}1&0&0\\0&1&0\end{pmatrix};\quad B=[0\quad 2\quad 0]$$

由(5-10)式，得

$$C_R^{-1}=C^T(CC^T)^{-1}$$

$$=\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\left(\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\right)^{-1}=\begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}$$

且满足

$$BC_R^{-1}C=B$$

再由公式(5-38)，得

$$\widetilde{A}^-=[C_R^{-1}\ \vdots\ 0]=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$$

于是有 $A^-=\widetilde{A}^-P=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}=\begin{pmatrix}0&0&1\\1&0&0\\0&0&0\end{pmatrix}$

4° 初等列变换法

设矩阵 $A=[a_{ij}]_{m\times n}$，且rank$A=r<\min(m,n)$，总可以作初等列变换将A化为下列分块矩阵

$$\widetilde{A}=[C\ \vdots\ B]$$

其中：C是列满秩的$m\times r$阶矩阵，B是$m\times(n-r)$阶矩阵，且满足$CC_L^{-1}B=B$，则可以验证

$$\widetilde{A}^-=\begin{pmatrix} C_L^{-1} \\ \hline 0 \end{pmatrix} \tag{5-40}$$

是$\widetilde{A}$的一个g逆，且具有反射性质。

由于对A作初等列变换，相当于右乘初等变换阵Q，即

$$\widetilde{A}=AQ$$

则有　　$A=\widetilde{A}Q^{-1}$

可以验证　　$A^-=(\widetilde{A}Q^{-1})^-=Q\widetilde{A}^-$　　(5-41)

显然，A^-是A的一个g逆，且具有反射性质。

〔**例8**〕 设

$$A=\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

求广义逆矩阵A^-。

解 因为$\operatorname{rank}A=2$，对A作初等列变换，得

$$\widetilde{A}=\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{array}\right);\quad Q=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$C=\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix};\quad B=\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

由(5-14)式，得

$$C_L^{-1}=(C^{T}C)^{-1}C^{T}=\left(\begin{pmatrix}1&0&0\\0&1&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\0&1\end{pmatrix}\right)^{-1}\begin{pmatrix}1&0&0\\0&1&1\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\0&1/2&1/2\end{pmatrix}$$

且满足 $CC_L^{-1}B=B$

再由公式(5-40)，得

$$\widetilde{A}^{-}=\begin{pmatrix}C_L^{-1}\\ \hline 0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1/2&1/2\\0&0&0\end{pmatrix}$$

于是，得到

$$A^{-}=Q\widetilde{A}^{-}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1/2&1/2\\1&0&0\end{pmatrix}$$

$$=\begin{pmatrix}0&0&0\\0&1/2&1/2\\1&0&0\end{pmatrix}$$

第二节 应用广义逆A^{-}解线性方程组

一、相容线性方程组的一般解

定理1 设矩阵$A=[a_{ij}]_{m\times n}$，如果线性方程组

$$AX=B \tag{5-42}$$

是相容的，则 $X=A^{-}B$

是线性方程组(5-42)的一个特解。

$$X=A^{-}B+(I-A^{-}A)C \tag{5-43}$$

是线性方程组(5-42)的一般解。

其中：A^-是A的任一个g逆，C是和X同维数的任意向量。

证明 因为$AX=B$相容，故必存在一个n维向量W，使得

$$AW=B$$

又因A^-是A的一个g逆，有

$$AA^-A=A \quad AA^-AW=AW$$

于是，有$AA^-B=B$

故 $X=A^-B$

即是方程组(5-42)的一个特解。

再证$X=(I-A^-A)C$是$AX=0$的一般解(即通解)。

设$\mu(A)$表示A的列向量所生成的m维线性空间的子空间，因此由A的行向量所生成的n维线性空间的子空间，可表示为$\mu(A^T)$。在n维线性空间中，和子空间$\mu(A^T)$垂直的一切向量组成的线性子空间叫做$\mu(A^T)$的正交补空间，记作$0(A^T)$。则有

$$\mu(A^T)\perp o(A^T)$$

显然，$AX=0$的所有解向量所组成的解空间是$\mu(A^T)$的正交补空间$o(A^T)$。下面证明$\mu(A^T)$的正交补空间是$(I-A^-A)C$。由于

$$A(I-A^-A)=0$$

有 $$A(I-A^-A)C=0$$

于是内积

$$(D, A(I-A^-A)C)=0 \quad (D\text{为任意}m\text{维向量})$$

因此，有

$$D^TA(I-A^-A)C=0$$

或 $$(A^TD)^T(I-A^-A)C=0$$

写成内积形式，有

$$(A^TD,\ (I-A^-A)C)=0$$

可见，$(I-A^-A)C$与A^TD垂直，而$A^TD\in\mu(A^T)$，又由于C是任意n维向量，所以$(I-A^-A)C$就构成了$\mu(A^T)$的正交补空间。因为在n维空间中$\mu(A^T)$的正交补空间是唯一的，所以$o(A^T)=(I-A^-A)C$。故$X=(I-A^-A)C$是齐次方程组$AX=0$的一般解。

非齐次方程组 $AX=B$的一般解，是它的任一特解及其对应齐次方程$AX=0$的一般解之和。

故有

$$X=A^-B+(I-A^-A)C \qquad \text{证毕。}$$

这一定理说明：一个相容线性方程组$AX=B$的系数矩阵A无论是方阵还是长矩阵，满秩的还是降秩的，都有一个标准的求解方法，其表达式即为(5-43)式。这是一个对线性方程组理论的重大发展。

求解线性方程组$AX=B$，只需求得A的一个g逆A^-，就可由(5-43)式，得到其通解。因此，求A^-就是求解相容线性方程组的一个重要步骤。求A^-的各种方法，前节已作出介绍。

(5-43)式又是相容方程组$AX=B$的通解的一种表达形式，此外还有一种表达形式。

定理2 相容非齐次线性方程组$AX=B$的一般解是

$$X=GB \qquad (5\text{-}44)$$

其中 $G=A^-+U-A^-AUAA^-$

或 $G=A^-+V(I-AA^-)+(I-A^-A)W$

U、V、W是任意适当阶矩阵。

证明 上面G的两个表达式就是本章第一节定理2中的

(5-23)式和(5-24)式，二者是等价的，故只要用其中之一来证明本定理即可。现用后者，即

$$G=A^{-}+V(I-AA^{-})+(I-A^{-}A)W$$

来证明。

由第一节定理2知，G为A的g逆，故$X=GB$为方程$AX=B$之解。

又若取$V=0$，$WB=C$，即得

$$X=GB=A^{-}B+(I-A^{-}A)C$$

由公式(5-43)可知，上式就是方程组$AX=B$的一般解。

二、相容线性方程组的最小范数解

相容线性方程组$AX=B$的所有解，可表示为$X=GB$，其中G是A的g逆。一般地，g逆有无穷多个，所以解向量X也有无穷多个。有必要讨论：是否存在与B无关的某些特殊g逆，使GB和其它的解相比较，具有最小范数(指向量的2-范数)，即

$$\|GB\|\leqslant\|X\| \tag{5-45}$$

其中：X是$AX=B$的解。

对此，我们有下面的定理

定理3 设G是A的g逆，GB在相容线性方程$AX=B$的一切解中具有最小范数的充要条件是：

$$AGA=A \quad 及(GA)^{\mathrm{T}}=GA \tag{5-46}$$

证明 因为$AX=B$的通解是$GB+(I-GA)C$，C是任意向量，若GB具有最小范数，则对C和一切与A构成相容方程的向量B，有

$$\|GB\|\leqslant\|GB+(I-GA)C\| \tag{5-47}$$

或 $$\|GAD\|\leqslant\|GAD+(I-GA)C\|$$

其中 $$B=AD$$

因此，

$$(GAD,GAD)^{\frac{1}{2}}\leqslant(GAD+(I-GA)C,GAD+(I-GA)C)^{\frac{1}{2}}$$

或$(GAD,\ GAD)\leqslant(GAD,\ GAD)+2(GAD,\ (I-GA)C)$
$+((I-GA)C,\ (I-GA)C)$

亦即 $0\leqslant((I-GA)C,\ (I-GA)C)+2(D,\ (GA)^{T}(I-GA)C)$由于上式右边第一项是向量$(I-GA)C$的范数平方，恒大于或等于零。于是以上不等式成立的充要条件是：

$$(GA)^{T}(I-GA)=0 \qquad (\text{因为}D,\ C\text{是任意的})$$

从而 $(GA)^{T}=(GA)^{T}GA$ 两边转置，得

$$GA=(GA)^{T}GA$$

比较两式，有 $(GA)^{T}=GA$

这就是定理的必要条件。

充分条件易于证明，请读者自己证。

我们把$X=GB$中具有最小范数的G称为**最小范数g逆**，常用A_{m}^{-}表示。

定理3说明：A_{m}^{-}是A的g逆受$(GA)^{T}=GA$限制的一个子集，最小范数g逆并不唯一，但相容方程组的最小范数解是唯一的。

定理4 相容线性方程组$AX=B$，具有唯一的最小范数解。

证明 需要先证明公式(5-46)与下式等价。

$$GAA^{T}=A^{T} \qquad (5\text{-}48)$$

两边右乘G^{T}，得

$$GAA^{T}G^{T}=A^{T}G^{T}$$

即 $GA(GA)^{T}=(GA)^{T}$ 两边转置

可得 $GA(GA)^{T}=GA$

则有 $(GA)^{T}=GA$ 将此式代入(5-48)式

得　　$(GA)^{T}A^{T}=A^{T}$　　两边转置

则有　$AGA=A$

反之，亦可由(5-46)式推出(5-48)式

此时，将$(GA)^{T}=GA$代入$AGA=A$，有

$$A(GA)^{T}=A \quad \text{两边转置，得}$$

$$GAA^{T}=A^{T}$$

故　(5-46)式和(5-48)式等价关系得以证明。

现在再证明定理4。

设G_1，G_2是两个不同的最小范数g逆，由(5-48)式，有

$$G_1AA^{T}=A^{T}, \qquad G_2AA^{T}=A^{T}$$

于是　$G_1AA^{T}=G_2AA^{T}$

即　$(G_1-G_2)AA^{T}=0$

又　$(G_1-G_2)AA^{T}(G_1^{T}-G_2^{T})=0$

有　$[(G_1-G_2)A][(G_1-G_2)A]^{T}=0$

仅当　$(G_1-G_2)A=0$时，上式方可成立。

故有

$$(G_1-G_2)AD=0 \qquad (D\text{为任意向量})$$

又因为$AD=B$，所以有　$G_1B=G_2B$

从而，对不同的最小范数g逆，按$X=GB$求得的最小范数解是唯一的得到证明。

此外，读者可以自行检证

$$A_m^{-}=A^{T}(AA^{T})^{-} \tag{5-49}$$

是A的一个最小范数g逆，而且是个反射g逆。最小范数g逆的一般表达式为：

$$A_m^{-}=A^{T}(AA^{T})^{-}+U(I-AA^{T}(AA^{T})^{-}) \tag{5-50}$$

其中U是具有适当阶的任意矩阵。

最小范数解的表达式为：

$$X = A_m^- B \tag{5-51}$$

〔例1〕 求方程组$AX=B$的最小范数解。

其中： $A=\begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$； $B=\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

解 由上节例1可知A是行满秩阵，因此AA^T是满秩方阵，AA^T的广义逆$(AA^T)^-$就是普通的逆矩阵$(AA^T)^{-1}$，所以由(5-49)式，得

$$A_m^- = A^T(AA^T)^- = A^T(AA^T)^{-1}$$

这正是行满秩阵的右逆，从而行满秩阵的右逆就是最小范数g逆。在上节例1中已求出A的右逆为

$$A_R^- = A_m^- = \frac{1}{14}\begin{pmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{pmatrix}$$

由(5-51)式得方程组的最小范数解

$$X = A_m^- B = \frac{1}{14}\begin{pmatrix} 5 & 4 \\ 6 & 2 \\ 3 & 8 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 13 \\ 10 \\ 19 \end{pmatrix} \tag{5-52}$$

由(5-43)式可得一般解为：

$$X = \frac{1}{14}\begin{pmatrix} 13+9c_1-6c_2-3c_3 \\ 10-6c_1+4c_2+2c_3 \\ 19-3c_1+2c_2+c_3 \end{pmatrix}$$

如令$c_1=c_2=0$，$c_3=1$，可求得一特解

$$X = \frac{1}{14}\begin{pmatrix} 10 \\ 12 \\ 20 \end{pmatrix} \tag{5-53}$$

分别对(5-52)式及(5-53)式求范数

$$\|X\| = (X^TX)^{\frac{1}{2}} = \frac{1}{14}\sqrt{13^2+10^2+19^2} = \frac{1}{14}\sqrt{630}$$

$$\|X\| = (X^TX)^{\frac{1}{2}} = \frac{1}{14}\sqrt{10^2+12^2+20^2} = \frac{1}{14}\sqrt{644}$$

显然，(5-52)式中X的范数小于(5-53)式中X的范数。

三、不相容方程组的最小二乘解

在实际问题的研究中，经常遇到不相容线性方程组的求解问题。但是从线性方程组理论上讲，这类方程组是无解的，因此有必要研究其最优近似解的问题。

对不相容的线性方程组 $AX=B$，如有

$$\|A\cdot\hat{X}-B\|\leqslant\|AX-B\| \tag{5-54}$$

成立，则$\hat{X}$是方程组$AX=B$在最小二乘意义下的**最优近似解**。因为和其它任何近似解X相比，$\hat{X}$所导致的误差平方和$\|A\cdot\hat{X}-B\|^2$是最小的。

定理5 设G是一个矩阵，对任意向量B，GB作为方程组$AX=B$的最小二乘解的充要条件是：

$$AGA=A \quad 和 \quad (AG)^T=AG \tag{5-55}$$

证明 先证必要条件：由假设有

$$\|AGB-B\|^2\leqslant\|AX-B\|^2$$

而
$$\begin{aligned}\|AX-B\|^2 &= \|AGB-B+AX-AGB\|^2\\ &= \|(AG-I)B+A(X-GB)\|^2\\ &= \|(AG-I)B+AC\|^2\end{aligned}$$

其中：$C=X-GB$

于是：$\|AGB-B\|^2=\|(AG-I)B\|^2\leqslant\|(AG-I)B+AC\|^2$

类似定理3的证明可知，上面不等式成立的充要条件是：

$$((AG-I)B,\ AC)=(B,\ (AG-I)^TAC)=0$$

可见，充要条件是： $(AG-I)^TA=0$

转置即得 $A^{T}AG=A^{T}$ (5-56)

两边右乘A，得 $A^{T}AGA=A^{T}A$

由第一节中性质3，则有$AGA=A$

再由(5-56)式两边左乘G^{T}，得

$$G^{T}A^{T}AG=G^{T}A^{T} \quad 即(AG)^{T}AG=(AG)^{T}$$

两边转置后，得

$$(AG)^{T}AG=AG$$

比较以上二式，可得 $(AG)^{T}=AG$

这就是必要条件。

关于充分条件，请读者自证。

应该指出，矛盾方程的最小二乘解导致的误差平方和$\|AX-B\|^{2}$是唯一的，而最小二乘解可以不唯一。

推论：设GB是一个最小二乘解，则矛盾方程组$AX=B$的最小二乘解的一般表达式为

$$\hat{X}=GB+(I-GA)C \tag{5-57}$$

C是任意向量。

证明 因为GB是矛盾方程的最小二乘解，因此$\|AGB-B\|^{2}$最小，从而

$$\|A(GB+(I-GA)C)-B\|^{2}=\|AGB-B\|^{2}$$

故说明(5-57)式也是矛盾方程的最小二乘解。由于C是任意向量，所以

$$\hat{X}=GB+(I-GA)C$$

是矛盾方程$AX=B$的最小二乘解的一般表达式。

又因使GB成为矛盾方程的最小二乘解的矩阵G应满足(5-55)式，所以G是一个g逆，称为**最小二乘g逆**，记为A_{l}^{-}。显然，A的最小二乘g逆是g逆在条件$(AG)^{T}=AG$限制下，g逆的一个子集。

根据第一节的性质2及(5-56)式，可以证明

$$(A^TA)^-A^T \tag{5-58}$$

及 $$(A^TA)^-A^T+(I-(A^TA)^-A^TA)U \tag{5-59}$$

是最小二乘g逆，其中U是任意矩阵，(5-59)式即为最小二乘g逆的一般表达式。

若对列满秩阵A，A^TA为满秩方阵，于是有

$$(A^TA)^-=(A^TA)^{-1}$$

因此，(5-58)式就是A的左逆。从而，列满秩阵的左逆就是最小二乘g逆。

〔**例2**〕 求矛盾方程组$AX=B$的最小二乘近似解：

其中： $A=\begin{pmatrix}1&2\\2&1\\1&1\end{pmatrix}$； $B=\begin{pmatrix}1\\0\\0\end{pmatrix}$； $X=\begin{pmatrix}X_1\\X_2\end{pmatrix}$

解 因A是列满秩阵，其左逆就是最小二乘g逆，在第一节例2中已求得A的左逆为：

$$A_L^{-1}=(A^TA)^{-1}A^T=\frac{1}{11}\begin{pmatrix}-4&7&1\\7&-4&1\end{pmatrix}=A_l^-$$

因此，矛盾方程组的最小二乘解

$$\hat{X}=A_l^-B=\frac{1}{11}\begin{pmatrix}-4&7&1\\7&-4&1\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\frac{1}{11}\begin{pmatrix}-4\\7\end{pmatrix}$$

从而可得

$$\hat{x}_1=\frac{-4}{11},\quad \hat{x}_2=\frac{7}{11}$$

由于 $$\|A\hat{X}-B\|^2=(A\hat{X}-B)^T(A\hat{X}-B)$$

$$=\left(\begin{pmatrix}1&2\\2&1\\1&1\end{pmatrix}\begin{pmatrix}\hat{x}_1\\\hat{x}_2\end{pmatrix}-\begin{pmatrix}1\\0\\0\end{pmatrix}\right)^T$$

$$\times\left(\begin{pmatrix}1&2\\2&1\\1&1\end{pmatrix}\begin{pmatrix}\hat{x}_1\\\hat{x}_2\end{pmatrix}-\begin{pmatrix}1\\0\\0\end{pmatrix}\right)$$

$$=(\hat{x}_1+2\hat{x}_2-1)^2+(2\hat{x}_1+\hat{x}_2)^2+(\hat{x}_1+\hat{x}_2)^2$$

代入$\hat{x}_1$，$\hat{x}_2$的值，即得

$$\|A\hat{X}-B\|^2=\frac{1}{11}$$

对于任意给出的矛盾方程组的近似解：

$$X=\begin{pmatrix}0\\0\end{pmatrix}\text{ 或 }X=\begin{pmatrix}1\\1\end{pmatrix}$$

其误差平方和

$$\|AX-B\|^2=1\text{或}\|AX-B\|^2=17$$

都大于$\frac{1}{11}$。

广义逆矩阵理论使矛盾方程组$AX=B$的求解问题归结为求取系数矩阵A的最小二乘g逆，无需采用先计算误差平方和，再利用极值条件进行求解的“古典法”，这就大大减少了计算工作量，并且更适用于电子计算机计算。求矛盾方程组的最小二乘解的问题，在最小二乘曲线拟合和多元线性回归分析中，是常见的。

以上所论述的是对任意矩阵A，由$AGA=A$定义的A的g逆A^-所进行的讨论，是一个最广义的逆矩阵概念。证明了A^-存在，但是不唯一。并指出了A^-的一个重要性质：$X=A^-B$是相容线性方程组$AX=B$的解；对A^-加以不同的限制，即可得到g逆的不同性质的子集。例如：

在$AGA=A$的基础上加以$GAG=G$的限制，得出的g逆的子集具有“反射”性质，称为反射g逆，记为A_r^-；

在$AGA=A$的基础上加以$(GA)^T=GA$的限制，得出的g逆的子集称为最小范数g逆，记为A_m^-。且具有重要性质：$X=A_m^-B$是相容线性方程组$AX=B$的最小范数解，而且是唯一的；

在$AGA=A$的基础上加以$(AG)^T=AG$的限制，得出的g逆的子集称为最小二乘g逆，记为A_l^-。且具有重要性质：$X=A_l^-B$是矛盾方程组$AX=B$的最小二乘解。

注意：A^-和A^-的子集A_r^-，A_m^-，A_l^-等一般都不是唯一的。

下面我们将讨论对g逆进一步加以限制条件，定义出新的广义逆A^+的问题，并讨论其性质及应用。

第三节 Moore-Penrose广义逆A⁺

一、广义逆A⁺的定义与性质

定义 设矩阵$A=[a_{ij}]_{m\times n}$，若存在矩阵$G=[g_{ij}]_{m\times n}$满足下列条件：

(1) $AGA=A$　　　　(2) $(GA)^T=GA$

(3) $GAG=G$　　　　(4) $(AG)^T=AG$

则称矩阵G为矩阵A的Moore-Penrose**广义逆**，记为A^+

定理1 对任意矩阵A，必存在唯一的Moore-Penrose广义逆A^+。

证明 首先证明广义逆A^+的存在性：

1°设A是行满秩阵，则其右逆$A_R^{-1}=A^T(AA^T)^{-1}$由第一节中(5-17)式(5-20)式知满足定义中全部条件，故右逆A_R^{-1}

就是矩阵A的Moore-Perose广义逆$A^+=A_R^{-1}$；

2°设A是列满秩阵，同理可知其左逆$A_L^{-1}=(A^TA)^{-1}A^T$也是矩阵A的Moore-Penrose广义逆$A^+=A_L^{-1}$；

3°设A是一般的$m\times n$矩阵，且$\text{rank}A=r<\min(m, n)$。由前面的讨论，总可作出$A$的满秩分解$A=CD$，其中$C=[c_{ij}]_{m\times r}$为列满秩阵，$D=[d_{ij}]_{r\times n}$为行满秩阵。再用左逆、右逆定义，$C_L^{-1}=(C^TC)^{-1}C^T$，$D_R^{-1}=D^T(DD^T)^{-1}$可得

$$G=D_R^{-1}C_L^{-1}=D^T(DD^T)^{-1}(C^TC)^{-1}C^T$$

容易验证G满足定义的全部条件，故G是A的Moore-Penrose广义逆$A^+=G$，

这就证明了对任意矩阵A，都存在Moore-Penrose广义逆A^+。

下面再证明唯一性。

需要说明：在上节定理4中已证明了定义中(1)和(2)条件与$GAA^T=A^T$等价。用同样方法也可以证明定义中条件(3)、(4)和$G=GG^TA^T$等价，于是Moore-Penrose广义逆A^+的定义可改为满足条件：

$$GAA^T=A^T\text{和}GG^TA^T=G \tag{5-60}$$

设G_1和G_2都是矩阵A的Moore-Penrose广义逆，或者是满足方程(5-60)的两个解，则

$$\begin{aligned}G_1&=G_1G_1^TA^T=G_1G_1^T(AG_2A)^T=G_1G_1^TA^TG_2^TA^T\\&=G_1(AG_1)^T(AG_2)^T=G_1AG_1AG_2=G_1AG_2\\&=G_1AG_2AG_2=G_1A(G_2A)^TG_2=G_1AA^TG_2^TG_2\\&=A^TG_2^TG_2=(G_2A)^TG_2=G_2AG_2=G_2\end{aligned}$$

这就证明了唯一性。　　证毕。

Moore-Penrose广义逆A^+具有以下主要性质：

性质1　广义逆A^+是一个g逆，因此$X=A^+B$是相容线性方程组$AX=B$的一个特解；$X=A^+B+(I-A^+A)C$是一般

解，其中C是任意向量。

性质2 $(A^+)^T=(A^T)^+$　　由g逆性质1即得

性质3 A^+是反射g逆，且

$(A^+)^+=A$　（由反射g逆定义及A^+的唯一性可得）

性质4 A^+是最小范数g逆，因此，$X=A^+B$是相容线性方程组$AX=B$的最小范数解（由A^+定义即得）。

性质5 A^+是最小二乘g逆，因此，$X=A^+B$是矛盾方程组$AX=B$的最小二乘解（由A^+定义即得）。

定理2 矛盾方程组$AX=B$的最小二乘最小范数解为：$\hat{X}=A^+B$

证明 上节定理5的推论已证明矛盾方程组的最小二乘解的一般表达式为：

$$GB+(I-GA)C \qquad (C为任意向量)$$

由于G是最小二乘g逆，应满足

$$AGA=A \qquad 及 \qquad (AG)^T=AG$$

若设G是最小二乘最小范数g逆，则GB就是矛盾方程组的最小二乘最小范数解，即对任意向量B和C有

$$\|GB\|\leqslant\|GB+(I-GA)C\|$$

成立。

由上节定理3的证明可知，上面不等式成立的充要条件是：

$$(GB,\ (I-GA)C)=(B,\ G^T(I-GA)C)=0$$

而上式得以成立的重要条件是：

$$G^T(I-GA)=0$$

即　$G^T=G^TGA$

用证明公式(5-46)与公式(5-48)等价的类似方法，可证得

$G^T=G^TGA$与$GAG=G$，$(GA)^T=GA$等价（留作习题）。因此最小二乘最小范数g逆是Moore-Penrose广义逆A^+。所以

$\hat{X}=A^+B$是矛盾方程组$AX=B$的最小二乘最小范数解。

二、广义逆A^+的计算

1°若A是满秩方阵，则$A^+=A^{-1}$

2°若A是对角方阵，即

$$A=\mathrm{diag}[d_1,\ d_2,\ \cdots,\ d_n]$$

其中：d_1，d_2，…，d_n均为实数，可验证

$$A^+=\mathrm{diag}[d_1^+,\ d_2^+,\ \cdots,\ d_n^+]$$

其中：
$$\begin{cases} d_i^+=0, & \text{当}d_i=0\text{时} \\ d_i^+=\dfrac{1}{d_i}, & \text{当}d_i\neq 0\text{时} \end{cases}$$

3°若A是行满秩阵，则

$$A^+=A^T(AA^T)^{-1}$$

需指出：n维非零行向量$a^T=[a_1,\ a_2,\ \cdots,\ a_n]$，可看作是行满秩的$1\times n$阶矩阵，因此行向量$a^T$的Moore-Penrose广义逆

$$(a^T)^+=(a^T)^T(a^T(a^T)^T)^{-1}=a(a^Ta)^{-1}$$

$$=\frac{a}{(a^Ta)}=\frac{a}{\sum_{i=1}^{n}a_i^2}$$

是一个n维列向量。

4°若A是列满秩阵，则

$$A^+=(A^TA)^{-1}A^T$$

对m维非零向量$b(b^T=[b_1\ \ b_2\ \ \cdots\ \ b_m])$可看作是列满秩的$m\times 1$阶矩阵，因此，列向量$b$的Moore-Penrose广义逆

$$b^+=(b^{\mathrm{T}}b)^{-1}b^{\mathrm{T}}=\frac{b^{\mathrm{T}}}{\sum_{i=1}^{m}b_i^2}$$

是一个m维的行向量。

5°若A是降秩的$m\times n$阶矩阵，可用满秩分解法求A^+（参看定理1的证明）。

6°初等变换法

设$A=[a_{ij}]_{m\times n}$且$\operatorname{rank}A=r<\min(m, n)$通过适当初等变换，可把$A$变为$\widetilde{A}$，即有

$$\widetilde{A}=PAQ$$

P，Q都是初等变换矩阵，

$$\widetilde{A}=\begin{pmatrix}\widetilde{A}_1 & \widetilde{A}_2\\ \widetilde{A}_3 & \widetilde{A}_4\end{pmatrix}$$

其中：$\widetilde{A}_1$是$r\times r$阶满秩方阵，且$\widetilde{A}_4=\widetilde{A}_3\widetilde{A}_1^{-1}\widetilde{A}_2$

利用此关系和Moore-Penrose广义逆定义，可直接验证

$$\widetilde{A}^+=\begin{pmatrix}\widetilde{A}_1^{\mathrm{T}}B\widetilde{A}_1^{\mathrm{T}} & \widetilde{A}_1^{\mathrm{T}}B\widetilde{A}_3^{\mathrm{T}}\\ \widetilde{A}_2^{\mathrm{T}}B\widetilde{A}_1^{\mathrm{T}} & \widetilde{A}_2^{\mathrm{T}}B\widetilde{A}_3^{\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\widetilde{A}_1^{\mathrm{T}}\\ \widetilde{A}_2^{\mathrm{T}}\end{pmatrix}B[\widetilde{A}_1^{\mathrm{T}}\quad \widetilde{A}_3^{\mathrm{T}}]\tag{5-61}$$

其中：$B=(\widetilde{A}_1\widetilde{A}_1^{\mathrm{T}}+\widetilde{A}_2\widetilde{A}_2^{\mathrm{T}})^{-1}\widetilde{A}_1(\widetilde{A}_1^{\mathrm{T}}\widetilde{A}_1+\widetilde{A}_3^{\mathrm{T}}\widetilde{A}_3)^{-1}$

〔**例1**〕 求矛盾方程组

$$\begin{cases}2x_3=1\\ x_1+x_2=1\\ x_3=1\\ x_1+x_2+x_3=1\end{cases}\tag{5-62}$$

的最小二乘最小范数解。

解 把(5-62)式写成矩阵方程$AX=B$

其中

$$A=\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix};\quad B=\begin{pmatrix}1\\1\\1\\1\end{pmatrix};\quad \wedge=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

$$\text{rank}A=2$$

对A作初等变换

$$\widetilde{A}=PAQ=\begin{pmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{pmatrix}\begin{pmatrix}0&0&2\\1&1&0\\0&0&1\\1&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$$

$$=\begin{pmatrix}1&1&1\\1&0&1\\0&1&0\\0&2&0\end{pmatrix}=\begin{pmatrix}\widetilde{A}_1&\widetilde{A}_2\\\widetilde{A}_3&\widetilde{A}_4\end{pmatrix}$$

其中：$\widetilde{A}_1=\begin{pmatrix}1&1\\1&0\end{pmatrix}$；$\widetilde{A}_2=\begin{pmatrix}1\\1\end{pmatrix}$；$\widetilde{A}_3=\begin{pmatrix}0&1\\0&2\end{pmatrix}$；

$$\widetilde{A}_4=\begin{pmatrix}0\\0\end{pmatrix}$$

利用(5-61)式，求A^+：

$$B=\left(\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}1\\1\end{pmatrix}[1\ \ 1]\right)^{-1}\begin{pmatrix}1&1\\1&0\end{pmatrix}$$

$$\times\left(\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix}+\begin{pmatrix}0&0\\1&2\end{pmatrix}\begin{pmatrix}0&1\\0&2\end{pmatrix}\right)^{-1}$$

$$=\begin{pmatrix}3&2\\2&2\end{pmatrix}^{-1}\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}2&1\\1&6\end{pmatrix}^{-1}=\frac{1}{22}\begin{pmatrix}2&-2\\-2&3\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix}$$

$$\times\begin{pmatrix}6&-1\\-1&2\end{pmatrix}=\frac{1}{22}\begin{pmatrix}-2&4\\8&-5\end{pmatrix}$$

于是

$$\widetilde{A}^{+}=\left[\begin{matrix}\frac{1}{22}\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}-2&4\\8&-5\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix} & \frac{1}{22}\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}-2&4\\8&-5\end{pmatrix}\begin{pmatrix}0&0\\1&2\end{pmatrix}\\ \frac{1}{22}[1\ \ 1]\begin{pmatrix}-2&4\\8&-5\end{pmatrix}\begin{pmatrix}1&1\\1&0\end{pmatrix} & \frac{1}{22}[1\ \ 1]\begin{pmatrix}-2&4\\8&-5\end{pmatrix}\begin{pmatrix}0&0\\1&2\end{pmatrix}\end{matrix}\right]$$

$$=\frac{1}{22}\begin{pmatrix}5&6&-1&-2\\2&-2&4&8\\5&6&-1&-2\end{pmatrix}$$

由于$\widetilde{A}=PAQ$，若将方程组$AX=B$写成

$$PAQQ^{-1}X=PB$$

并记$\widetilde{X}=Q^{-1}X$，$\widetilde{B}=PB$，则方程组$AX=B$就转化成$\widetilde{A}\ \widetilde{X}=\widetilde{B}$，其最小二乘最小范数解为：

$$\hat{\widetilde{X}}=\widetilde{A}^{+}\ \widetilde{B}$$

依此

$$\hat{\tilde{X}} = Q^{-1}\hat{X} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_3 \\ \hat{x}_2 \end{pmatrix}$$

其中：$\hat{X}$为原方程组$AX=B$的最小二乘最小范数解。由于

$$\tilde{B} = PB = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

则

$$\hat{\tilde{X}} = \tilde{A}^{+}\tilde{B} = \frac{1}{22}\begin{pmatrix} 5 & 6 & -1 & -2 \\ 2 & -2 & 4 & 8 \\ 5 & 6 & -1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$= \frac{1}{11}\begin{pmatrix} 4 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_3 \\ \hat{x}_2 \end{pmatrix}$$

可见

$$\hat{X} = \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{pmatrix} = \frac{1}{11}\begin{pmatrix} 4 \\ 4 \\ 6 \end{pmatrix}$$

这就是(5-62)式的最小二乘最小范数解（也称为最优近似解）。

从本例可以看到：由A^{+}阵求最优近似解比古典法来说，理论严谨、清晰，方法简单方便。

〔例2〕 求相容方程组

$$\begin{cases} \qquad\qquad 2x_3=2 \\ x_1+x_2 \qquad =0 \\ \qquad\qquad x_3=1 \\ x_1+x_2+x_3=1 \end{cases} \tag{5-63}$$

的最小范数解。

解 把(5-63)式写成矩阵方程

$$AX=B$$

其中：

$$A=\begin{pmatrix} 0 & 0 & 2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}; \quad B=\begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix}; \quad X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

实行和例 1 相同的初等变换后，得

$$\widetilde{A}\,\widetilde{X}=\widetilde{B}$$

其中：

$$\widetilde{A}=PAQ=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{pmatrix}; \quad \widetilde{X}=Q^{-1}X=\begin{pmatrix} x_1 \\ x_3 \\ x_2 \end{pmatrix};$$

$$\widetilde{B}=PB=\begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

由例 1 知

$$\tilde{A}^+ = \frac{1}{22}\begin{pmatrix} 5 & 6 & -1 & -2 \\ 2 & -2 & 4 & 8 \\ 5 & 6 & -1 & -2 \end{pmatrix}$$

于是

$$\tilde{A} = \tilde{A}^+\tilde{B} = \frac{1}{22}\begin{pmatrix} 5 & 6 & -1 & -2 \\ 2 & -2 & 4 & 8 \\ 5 & 6 & -1 & -2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 2 \end{pmatrix}$$

$$= \frac{1}{22}\begin{pmatrix} 0 \\ 22 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_3 \\ x_2 \end{pmatrix}$$

从而相容方程组的最小范数解为：

$x_1=0$， $x_2=0$， $x_3=1$

其范数是

$$\|X\| \sqrt{0^2+0^2+1^2}=1$$

而方程组的其它解，例如：$x_1=-2$，$x_2=2$，$x_3=1$其范数 $\|X\|=\sqrt{(-2)^2+2^2+1^2}=3\ (>1)$，显然大于上述解的范数。

此外，矩阵A恒可通过一系列初等变换变成：

$$\tilde{A}=PAQ=\begin{pmatrix} \tilde{A}_1 & 0 \\ 0 & 0 \end{pmatrix}$$

其中：$\tilde{A}_1$为$r\times r$阶满秩方阵。

此时，广义逆$\tilde{A}^+$就有以下公式

$$\tilde{A}^+ = \begin{pmatrix} \tilde{A}_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} \tag{5-64}$$

应注意到：由(5-61)式或(5-64)式所求得的$\tilde{A}^+$是矩阵$\tilde{A}$的

Moore-Penrose广义逆，而不是矩阵A的Moore-Penrose广义逆。在一些问题的讨论中(如例1，例2)只需求取$\widetilde{A}^+$。倘若需要求取A的广义逆$\widetilde{A}^+$时，则由$\widetilde{A}=PAQ$可验证

$$A^+=Q\widetilde{A}^+P \tag{5-65}$$

但是，其中初等变换阵P、Q应为正交矩阵，即

$$P^T=P^{-1},\quad Q^T=Q^{-1}$$

若P、Q不是正交矩阵，则(5-65)式并不成立。这时，A^+亦可采用满秩分解法或另外别的方法求取。

还应指出：矩阵A的广义逆A^-与A^+对初等变换阵P、Q的要求上是有所区别的。因为由

$\widetilde{A}=PAQ$　可以验证

$$A^-=Q\widetilde{A}^-P$$

而其中的初等变换阵P、Q并不一定要求是正交矩阵。这是由于按广义逆的定义A^-只需满足一个方程

$$AA^-A=A$$

而A^+则要满足四个方程

$$AA^+A=A \qquad (A^+A)^T=A^+A$$

$$A^+AA^+=A^+ \qquad (AA^+)^T=A^+$$

第四节　应用广义逆矩阵解各种矩阵方程

定理1　矩阵方程$AXB=C$　(5-66)

有解的充要条件是$AA^-CB^-B=C$　(5-67)

方程(5-66)的一般解为：

$$X=A^-CB^-+M-A^-AMBB^- \tag{5-68}$$

其中：M是与X同类型的任意矩阵。



证明　先证明条件的必要性：
若方程(5-66)有解，必有一矩阵X_0，满足

$$AX_0B\equiv C$$

于是有

$$AA^-CB^-B=AA^-AX_0BB^-B=AX_0B\equiv C$$

再证明条件的充分性：
若条件(5-67)式成立，则A^-CB^-是方程(5-66)的解。

还要证明(5-68)式是方程(5-66)的一般解只需证明方程(5-66)的任一解均可表示为(5-68)式形式。设X_0是方程(5-66)的任一解，只要取$M=X_0-A^-CB^-$，就可把X_0表示成(5-68)式形式。至此，定理证毕。

推论：矩阵方程(5-66)的几种特殊形式：

(1)　设$B=I$，则方程(5-66)变为：$AX=C$
其有解的充要条件相应地变为：$AA^-C=C$
因此，一般解也相应地变为：$X=A^-C+(I-A^-A)M$
倘若在方程$AX=C$中，C是向量，那么X也是向量，$AX=C$就是通常的线性方程组；$AA^-C=C$就是线性方程组相容的充要条件的另一种形式。

(2)　设$A=I$，则方程(5-66)变为：$XB=C$
其有解的充要条件相应变为：$CB^-B=C$
因此，一般解也相应变为：$X=CB^-+M(I-BB^-)$

(3)　设$B=A^T$，则方程(5-66)变为：$AXA^T=C$
其有解的充要条件相应变为：$(AA^-)C(AA^-)^T=C$
一般解相应地有

$$X=A^-C(A^-)^T+M-A^-AM(A^-A)^T$$

定理2　分别相容的二矩阵方程

$$AX=C \quad 及 \quad XB=D \tag{5-69}$$

有公共解的充要条件是

$$AD=CB \tag{5-70}$$

其公共解的一般形式应为

$$X=A^{-}C+DB^{-}-A^{-}ADB^{-}+(I-A^{-}A)M(I-BB^{-}) \tag{5-71}$$

其中M为任意矩阵。

证明 关于必要性：若公共解存在，则

$$AXB=CB=AD$$

关于充分性：设$AD=CB$，取

$$X=A^{-}C+DB^{-}-A^{-}ADB^{-} \tag{5-72}$$

因$AX=C$相容，则利用定理1的推论(1)，可以验证(5-72)式是$AX=C$的一个解。

把(5-72)式代入$XB=D$，则利用$AD=CB$有

$$\begin{aligned}XB&=A^{-}CB+DB^{-}B-A^{-}ADB^{-}B\\&=A^{-}CB+DB^{-}B-A^{-}CBB^{-}B\\&=A^{-}CB+DB^{-}B-A^{-}CB=DB^{-}B\end{aligned}$$

因$XB=D$相容，再利用定理1的推论(2)，得

$$XB=DB^{-}B=D$$

于是(5-72)式是方程(5-69)的公共解。充分性得以证明。

对于(5-71)式是(5-69)式的通解问题，只需证明

$$X=(I-A^{-}A)M(I-BB^{-}) \tag{5-73}$$

是齐次矩阵方程

$$\begin{cases}AX=0\\XB=0\end{cases} \tag{5-74}$$

的通解即可。

首先容易直接验证(5-73)式是方程(5-74)的解。当然，还应

该证明方程(5-74)的任一解均可表示为(5-73)式。为此，设X_0是方程(5-74)的任一解，对X_0进行满秩分解：$X_0=UV$，其中U是列满秩阵，V是行满秩阵，其秩等于X_0的秩，即$\text{rank}U=\text{rank}V=\text{rnak}X_0$，由$AX_0=0$，可推知$AU=0$，根据定理1的推论(1)知齐次矩阵方程

$$AU=0$$

的一般解为：

$$U=(I-A^-A)M_1$$

其中：M_1是任意矩阵。又由$X_0B=0$，可推知$VB=0$，再根据定理1的推论(2)知齐矩阵方程

$$VB=0$$

的一般解为：$V=M_2(I-BB^-)$

其中：M_2是任意矩阵。

由此可知，方程(5-74)的任意解X_0均可表示为

$$X_0=UV=(I-A^-A)M_1M_2(I-BB^-)$$

$$=(I-A^-A)M(I-BB^-)$$ 证毕。

下面我们利用各种广义逆来讨论矩阵方程的解。

定理3 矩阵方程$XBX=X$ (5-75)

的一般解为：$X=M(NBM)_r^-N$ (5-76)

且 $\text{rank}X=\text{rank}(NBM)$ (5-77)

其中：B是$m\times n$阶矩阵，M和N分别是$n\times p$阶和$q\times m$阶的任意矩阵，$(NBM)_r^-$是(NBM)的反射g逆。

根据本章第一节中定义4知，方程(5-75)

和 $X^-=B$ (5-78)

等价。即是求解方程(5-75)就是要找出一个矩阵其g逆等于已给矩阵B。

证明 显然，(5-76)式是方程(5-75)的解。只需证明方

程(5-75)的任一解都包含在(5-76)式中。为此，设X_0是方程(5-75)的解，则有

$$X_0BX_0=X_0 \tag{5-79}$$

两端左乘B，得

$$(BX_0)(BX_0)=BX_0$$

由此知BX_0是幂等矩阵⑤。于是可以推出

$$(BX_0)(BX_0)(BX_0)=BX_0$$

因此，BX_0是自身的一个反射g逆，即$BX_0=(BX_0)_r^-$今在(5-76)式中取$M=X_0$，$N=I$，得

$$X=X_0(BX_0)_r^-$$

以$(BX_0)_r^-=BX_0$代入，则(5-79)式即成为：

$$X=X_0(BX_0)=X_0$$

这就证明了方程(5-75)的任一解都包含在(5-76)式中。故(5-76)式是方程(5-75)的一般解。

从(5-76)式，可得$\mathrm{rank}X\leqslant\mathrm{rank}(NBM)_r^-$利用广义逆$A^-$的性质4及反射g逆的定义，可证明

$$\mathrm{rank}(NBM)_r^-=\mathrm{rank}(NBM)$$

又 $NBM=NBM(NBM)_r^-NBM=NBXBM$，则有

$$\mathrm{rank}(NBM)\leqslant\mathrm{rank}X$$

因此，可得

$$\mathrm{rank}X=\mathrm{rank}(NBM)$$ 证毕。

定理4 矩阵方程

$$XBX=X,\quad (BX)^T=BX \tag{5-80}$$

的一般解为

$$X=M(NBM)_{l,r}^-N \tag{5-81}$$

---

⑤矩阵自乘仍等于自身者，称为幂等矩阵

其中：$(NBM)_{lr}^{-}$是(NBM)的反射最小二乘g逆。M是任意矩阵，而N是满足$N^{T}N=I_{m}$的任意矩阵。

须指出：方程(5-80)和

$$X_{m}^{-}=B \tag{5-82}$$

等价，即方程(5-80)的解矩阵X的最小范数g逆等于已给矩阵B。

证明 容易验证(5-81)式满足(5-80)中第一个方程。今证其也满足第二个方程：由(5-81)式有

$$BX=BM(NBM)_{lr}^{-}N=N^{T}NBM(NBM)_{lr}^{-}N$$
$$=N^{T}DN$$

其中：$D=NBM(NBM)_{lr}^{-}$，根据最小二乘g逆的定义知$D^{T}=D$。于是有

$$(BX)^{T}=(N^{T}DN)^{T}=N^{T}D^{T}N=N^{T}DN=BX$$

这就证明了(5-81)式满足(5-80)中的两个方程。

还需证明(5-81)式是方程(5-80)的一般解。为此，设X_0是方程(5-80)的解，则有

$$X_0BX_0=X_0 \qquad (BX_0)^{T}=BX_0$$

由第一式可得$(BX_0)(BX_0)=BX_0$，说明BX_0为幂等矩阵，从而可推出$(BX_0)(BX_0)(BX_0)=BX_0$，可见BX_0又是其自身的一个反射g逆。又由上面第二式，根据BX_0为幂等矩阵，可得

$$(BX_0BX_0)^{T}=BX_0BX_0$$

这说明BX_0又是其自身的一个最小二乘g逆，由此证得BX_0是其自身的一个反射最小二乘g逆，即$(BX_0)_{lr}^{-}=BX_0$

今在(5-81)式中令$M=X_0$，$N=I$，得

$$X=X_0(BX_0)_{lr}^{-}$$

以$(BX_0)_{lr}^{-}=BX_0$代入，则(5-81)式即成为

$$X=X_0(BX_0)=X_0$$

这就证明了方程(5-80)的任一解都包含在(5-81)式中，故(5-81)式是方程(5-80)的一般解。

定理5 矩阵方程

$$XBX=X, \quad (XB)^{T}=XB \tag{5-83}$$

与矩阵方程

$$X_{l}^{-}=B \tag{5-84}$$

等价。其一般解为：

$$X=M(NBM)_{mr}^{-}N \tag{5-85}$$

其中：N是任意矩阵，M是满足$MM^{T}=I$的任意矩阵，$(NBM)_{mr}^{-}$是(NBM)的反射最小范数g逆。

证明： 仿定理4的证明(过程略)。

定理6 矩阵方阵

$$XBX=X, \quad BXB=B \tag{5-86}$$

与矩阵方程$X_{r}^{-}=B$　　(5-87)

等价。其一般解为$X=B_{r}^{-}$　　(5-88)

证明 由于(5-86)式是反射g逆的定义式，是完全对称的，即X是B的反射g逆。而B也是X的反射g逆。

由定理3知，方程$XBX=X$的一般解是

$$X=M(NBM)_{r}^{-}N$$

由于X还须同时满足$BXB=B$

故方程(5-86)的解只应为$X=B_{r}^{-}$

即是在$X=M(NBM)_{r}^{-}N$中M、N均为单位阵时的特殊情形。

定理7 矩阵方程

$$XBX=X, \quad BXB=B, \quad (BX)^{T}=BX \tag{5-89}$$

等价于矩阵方程$X_{mr}^{-}=B$　　(5-90)

其一般解是$X=B_{lr}^{-}$　　(5-91)

就是说B的反射最小二乘g逆是方程(5-89)或方程(5-90)

的解。

证明 方程(5-89)比方程(5-80)多一个限制条件 $BXB=B$，而方程(5-80)的一般解是

$$X=M(NBM)_{l\,r}^{-}N$$

由于X还须同时满足条件$BXB=B$，故方程(5-89)之解只应为

$$X=B_{l\,r}^{-},$$

即是在$X=M(NBM)_{l\,r}^{-}$中，M、N均为单位阵时的特殊情形。

定理8 矩阵方程

$$XBX=X,\quad BXB=B,\quad (XB)^{T}=XB \tag{5-92}$$

等价于矩阵方程$X_{l\,r}^{-}=B$ (5-93)

其一般解是$X=B_{mr}^{-}$ (5-94)

证明 类似于定理7的证明(过程略)。

方程(5-92)的一般解

$$X=B_{mr}^{-}$$

即在(5-85)式中，M、N均为单位阵的特殊情形。

定理9 矩阵方程

$$XBX=X,\quad BXB=B,\quad (XB)^{T}=XB,\quad (BX)^{T}=BX \tag{5-95}$$

等价于矩阵方程$X^{+}=B$ (5-96)

其唯一解为$X=B^{+}$ (5-97)

证明 由上节定义及定理1即得。

定理10 矩阵方程

$$XBX=0 \tag{5-98}$$

的一般解是 $X=YC$ (5-99)

其中：B是$m\times n$阶矩阵，C是$p\times m$阶任意矩阵，p也是任意的，Y是方程

$$CBY=0 \qquad (5\text{-}100)$$

的任一个$n\times p$阶矩阵解。

证明 $X=YC$显然是方程(5-98)的解。现只需证明(5-98)式的任一解X均可表示为(5-99)式的形式。将X作满秩分解，令$X=DC$，而X又是方程(5-98)的解，因此有$DCBDC=0$，于是得到$CBD=0$，可见D就是方程(5-100)的解，故解X可表成(5-99)式的形式。

推论 矩阵方程

$$XBX=0,\quad WBX=0 \qquad (5\text{-}101)$$

的一般解为 $X=YC$ (5-102)

其中C是任意$p\times m$阶矩阵，p是任意的，Y是方程

$$\begin{bmatrix} C \\ W \end{bmatrix} BY=0 \qquad (5\text{-}103)$$

的任一解。

定理11 不相容矩阵方程

$$AXB=C \qquad (5\text{-}104)$$

的最小范数最小二乘解是

$$\hat{X}=A^{+}CB^{+} \qquad (5\text{-}105)$$

其中A、B、C都是已给矩阵，A^{+}、B^{+}分别是A，B的Moore-Penrose广义逆。

前面定理1给出了矩阵方程$AXB=C$的相容条件和在此条件下的一般解。定理11给出了矩阵方程$AXB=C$不相容时的最小范数最小二乘解$\hat{X}$，就是在使得$\|AXB-C\|^2$最小的一切矩阵X中，$\hat{X}$矩阵的范数$\|\hat{X}\|$是最小的，即 $\|\hat{X}\|\leqslant\|X\|$。

习　题　五

1. 证明

(1)　$\operatorname{rank}(A^{-}A)=\operatorname{rank}A$

(2)　$\operatorname{rank}(A_r^{-})=\operatorname{rank}A$

中A^{-}，A_r^{-}分别是矩阵A的g逆和反射g逆。

2. (1)　设A是$m\times n$阶矩阵，且$\operatorname{rank}A=r\leqslant\min(m,n)$，存在$m$阶矩阵$P$和$n$阶矩阵$Q$，使$B=PAQ$（$P$、$Q$均为非奇异矩阵）。

证明：$B^{-1}=Q^{-1}A^{-}P^{-1}$是矩阵B的g逆。

(2) 设A是$m\times n$阶矩阵，C、D都是矩阵A的g逆。

证明：$A_r^{-}=CAD$是矩阵A的反射g逆。

3. 求下列矩阵

$$A_1=\begin{pmatrix}1&0&2\\0&1&0\\1&0&2\\1&0&2\end{pmatrix},\quad A_2=\begin{pmatrix}2&1&0&1\\1&0&1&1\\1&0&1&1\end{pmatrix}$$

的g逆A_1^{-}和A_2^{-}。

并分别求矩阵方程：

$$A_1X_1=B_1\quad 和\quad A_2X_2=B_2$$

的一般解。

其中：$B_1=[1,0,1,1]^{T}$；$B_2=[2,1,1]^{T}$

4. 验证$A^{T}(AA^{T})^{-}$是A的一个最小范数g逆。

5. 求矩阵方程

$$AX=B$$

的最小范数解与A的最小范数g逆。

其中：

$$A=\begin{pmatrix}1&0&3\\2&3&0\\1&1&1\end{pmatrix},\quad B=\begin{pmatrix}3\\0\\1\end{pmatrix}$$

6．设G是任意$n\times m$阶矩阵，A是$m\times n$阶矩阵。

证明：(1) $G=GG^TA^T$与$GAG=G$，$(AG)^T=AG$等价；

(2) $A^TAG=A^T$与$AGA=A$，$(AG)^T=AG$等价；

(3) $G^TGA=G^T$与$GAG=G$，$(GA)^T=GA$等价。

7．用不同的方法求下列矩阵的Moore-Penrose广义逆A^+

(1) $A=\begin{pmatrix}1 & -1\\ 2 & -2\\ 4 & -4\end{pmatrix}$　　(2) $A=\begin{pmatrix}1 & 2 & 1\\ 0 & 1 & 1\end{pmatrix}$

(3) $A=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 1 & 0 & 0\\ 2 & 1 & -1\end{pmatrix}$

(4) $A=\begin{pmatrix}1 & 1 & 1\\ 2 & 0 & 0\\ 0 & 1 & 1\\ 1 & 0 & 0\end{pmatrix}$

8．求相容方程组$AX=B$的通解和最小范数解。

其中：

$A=\begin{pmatrix}1 & 2\\ 0 & 0\\ 2 & 4\end{pmatrix}$；　$B=\begin{pmatrix}1\\ 0\\ 2\end{pmatrix}$

9．求矛盾方程组$AX=B$的最小二乘最小范数解，

其中：

$A=\begin{pmatrix}1 & 2\\ 0 & 0\\ 2 & 4\end{pmatrix}$；　$B=\begin{pmatrix}1\\ 1\\ 2\end{pmatrix}$

习 题 答 案

习题一

1. (1) 是； (2) 不是。

2. (1) 不是；是；

   (2) 是。

5. $e=\frac{1}{\sqrt{26}}(4,\ 0,\ 1,\ -3)$

10. 不变因子：$d_1=1$，$d_2=\lambda-1$，$d_3=(\lambda+1)(\lambda-1)^2$

    初等因子：$\lambda-1$，$\lambda+1$，$(\lambda-1)^2$

11. $A\sim J=\begin{pmatrix}1 & & \\ & -i & \\ & & i\end{pmatrix}$；$B\sim J=\begin{pmatrix}-1 & & & & & \\ & -1 & & & & \\ & & 2 & & & \\ & & & 1 & 1 & \\ & & & & 1 & \\ & & & & & 1\end{pmatrix}$

12. $J=\begin{pmatrix}1 & & \\ & 1 & 1 \\ & & 1\end{pmatrix}$，$P=\begin{pmatrix}1 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & -1 & -1\end{pmatrix}$ (P不唯一)

习题二

2. (1) $(A-I)^{-1}=\frac{1}{3}\begin{pmatrix}2 & 1 \\ -1 & 1\end{pmatrix}$

   (2) $f(B)=29A^2-24A+6I=\begin{pmatrix}301 & -744 & 126 \\ 92 & -231 & 34 \\ 212 & 520 & 105\end{pmatrix}$

3. $\varphi_m^A(\lambda)=(\lambda+9)(\lambda-9)$

   $\varphi_m^B(\lambda)=a_1^2+a_2^2+a_3^2+(\lambda-a_0)^2$

习题三

1. $$\frac{d^2}{dt^2}A(t)=\begin{pmatrix} -\sin t & -\cos t & 0 \\ \frac{2\sin t-2t\cos t-t^2\sin t}{t^3} & e^t & 2 \\ 0 & 0 & 6t \end{pmatrix}$$

$$\frac{d}{dt}\det A(t)=e^t[3t^2\sin t+t^3(\sin t+\cos t)-t-1]+t(2\cos t-t\sin t)-t(\sin 2t+t\cos 2t)$$

$$\det\frac{d}{dt}A(t)=3t^2e^t\cos t+3\sin(t\cos t-\sin t)$$

$$\frac{d}{dt}A^{-1}(t)=-\frac{1}{|A(t)|^2}A^*(t)\ \frac{d}{dt}A(t)A^*(t)$$

其中：$\det A(t)=|A(t)|=t^3e^t\sin t+t^2\cos t-te^t-t^2\cos t\sin t$

$$A^*(t)=\begin{pmatrix} t^3e^t & -t^3\cos t & t(t\cos t-e^t) \\ t^2(1-\sin t) & t(t^2\sin t-1) & \sin t(1-t^2) \\ -e^t & \cos t & e^t\sin t-\frac{\sin 2t}{2t} \end{pmatrix}$$

5. $$e^A=\frac{1}{7}\begin{pmatrix} 3e^5+4e^{-2} & 4e^5-4e^{-2} \\ 3e^5-3e^{-2} & 4e^5+3e^{-2} \end{pmatrix}$$

$$e^B=\frac{1}{3}[B+(\frac{1}{2}-\frac{\sqrt{3}}{2}i)I][B+(\frac{1}{2}+\frac{\sqrt{3}}{2}i)I]e$$
$$+\frac{\sqrt{3}i-1}{6}(B-I)[B+\frac{1+\sqrt{3}i}{2}I]e^{-\frac{1}{2}+\frac{\sqrt{3}i}{2}}$$
$$-\frac{1+\sqrt{3}i}{6}(B-I)[B+(\frac{1}{2}-\frac{\sqrt{3}}{2}i)I]e^{-\frac{1+\sqrt{3}i}{2}}$$

$$\sin B=P\begin{pmatrix} \sin 1 & & \\ & -\sin\frac{1+\sqrt{3}i}{2} & \\ & & -\sin\frac{1-\sqrt{3}i}{2} \end{pmatrix}P^{-1}$$

$$\cos B = P\begin{pmatrix} \cos 1 & & \\ & \cos\frac{1+\sqrt{3}i}{2} & \\ & & \cos\frac{1-\sqrt{3}i}{2} \end{pmatrix} P^{-1}$$

其中：

$$P = \begin{pmatrix} 1 & -\frac{1+\sqrt{3}i}{2} & \frac{-1+\sqrt{3}i}{2} \\ 1 & \frac{-1+\sqrt{3}i}{2} & -\frac{1+\sqrt{3}i}{2} \\ 1 & 1 & 1 \end{pmatrix},$$

$$P^{-1} = \frac{1}{3\sqrt{3}i}\begin{pmatrix} \sqrt{3}i & \sqrt{3}i & \sqrt{3}i \\ \frac{-3-\sqrt{3}i}{2} & \frac{3-\sqrt{3}i}{2} & \sqrt{3}i \\ \frac{3-\sqrt{3}i}{2} & \frac{-3-\sqrt{3}i}{2} & \sqrt{3}i \end{pmatrix}$$

6. $\mathrm{Sin}^{-1}(\frac{1}{4})A = (\frac{\pi}{6} - \frac{\sqrt{3}}{3})I + \frac{\sqrt{3}}{6}A$

8. $\begin{pmatrix} -184 & 189 \\ -189 & 194 \end{pmatrix}$

9. $\frac{5I-A}{4} + \frac{A-I}{4}5^{1000}$

10. $e^{At} = \begin{pmatrix} 2e^{t}-e^{2t} & 0 & 2e^{t}-2e^{2t} \\ 0 & e^{t} & 0 \\ 2e^{2t}-e^{t} & 0 & 2e^{2t}-e^{t} \end{pmatrix}$

$$e^{Bt} = \begin{pmatrix} e^{3t} & 0 & 0 & 0 \\ 0 & e^{-2t} & te^{-2t} & \frac{t^2}{2}e^{-2t} \\ 0 & 0 & e^{-2t} & te^{-2t} \\ 0 & 0 & 0 & e^{-2t} \end{pmatrix}$$

$$e^{Ct}=\begin{pmatrix} 2e^{-t}-e^{-2t} & e^{-t}-e^{-2t} \\ -2(e^{-t}-e^{-2t}) & -e^{-t}+2e^{-2t} \end{pmatrix}$$

习题四

1. (1)
$$\begin{cases} x_1(t)=2(c_1-c_3)-(c_1-2c_3)e^t+(c_1+c_2-c_3)te^t \\ x_2(t)=c_3-c_1+(c_1+c_2-c_3)e^t \\ x_3(t)=c_1-c_3+(-c_1+2c_3)e^t+(c_1+c_2-c_3)te^t \end{cases}$$

(2)
$$\begin{cases} x_1(t)=c_1\cos t+c_2\sin t \\ x_2(t)=-c_1\sin t+c_2\cos t \end{cases}$$

(3)
$$\begin{cases} s_1(t)=(c_3-c_3t-c_2)e^t \\ s_2(t)=(c_3-2c_3t-2c_2)e^t \\ s_3(t)=c_1e^{2t}+(c_2+c_3t)e^t \end{cases}$$

2.

(1)
$$\begin{cases} x_1(t)=(\frac{1}{2}-\frac{\sqrt{2}}{4})e^{(\sqrt{2}-4)}+(\frac{1}{2}+\frac{\sqrt{2}}{4})e^{-(4+\sqrt{2})t} \\ x_2(t)=\frac{\sqrt{2}}{4}(e^{(\sqrt{2}-4)t}-e^{-(\sqrt{2}+4)t}) \end{cases}$$

(2) (a)
$$\begin{cases} x_1(t)=\frac{1}{3}[(c_1+c_2+\frac{1}{2})e^{5t}+(2c_1-c_2-\frac{1}{2})e^{-t}]-2te^{-t} \\ x_2(t)=\frac{1}{3}\{[2(c_1+c_2)+1]e^{5t}-(2c_1-c_2+1)e^{-t}\}+2te^{-t} \end{cases}$$

(b)
$$\begin{cases} x_1(t)=e^{-t}-4e^{-2t}+\frac{11}{3}e^{-3t}+\frac{1}{[illegible]} \\ x_2(t)=5e^{-t}-16e^{-2t}+11e^{-3t} \\ x_3(t)=6e^{-t}-12e^{-2t}+\frac{22}{3}e^{-3t}-\frac{7}{3} \end{cases}$$

(3)
$$\begin{cases} x_1(t)=1-t+\frac{3}{2}t^2-\frac{1}{2}t^3+\frac{1}{8}t^4-\cdots \\ x_2(t)=-1+2t-\frac{3}{2}t^2+t^3-\frac{1}{8}t^4+\cdots \end{cases}$$

3.(1) $\begin{cases} x_1(t)=1-t+\frac{3}{2}t^2-\frac{1}{3}t^3+\frac{1}{8}t^4-\frac{53}{120}t^5+\cdots \\ x_2(t)=-1+2t-t^2+t^3+\frac{1}{12}t^4-\frac{1}{24}t^6+\cdots \end{cases}$

(2) $\begin{cases} x_1(t)=2-e^t-\frac{1}{12}t^4+\frac{1}{20}t^5+\frac{1}{224}t^8+\cdots \\ x_2(t)=3-4e^t+3te^t-t^2e^t+\frac{1}{4}t^4+\frac{1}{18}t^6+\cdots \end{cases}$

习题五

3. $A_1^- = \frac{1}{15}\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 15 & 0 & 0 \\ 2 & 0 & 2 & 2 \end{pmatrix}$; $A_2^- = \frac{1}{6}\begin{pmatrix} 2 & 0 & 0 \\ 2 & -1 & -1 \\ -2 & 2 & 2 \\ 0 & 1 & 1 \end{pmatrix}$

$X_1=\frac{1}{5}\begin{pmatrix} 1+4c_1-2c_3 \\ 2-2c_1+c_3 \end{pmatrix}$; $X_2=\frac{1}{3}\begin{pmatrix} 2+c_1-c_2-c_4 \\ 1-c_1+2c_2+c_3 \\ c_2+c_3-c_4 \\ 1-c_1-c_3+2c_4 \end{pmatrix}$

5. $A^-_m=\frac{1}{462}\begin{pmatrix} 33 & 66 & 0 \\ -22 & 110 & 0 \\ 143 & -22 & 0 \end{pmatrix}$; $X=\frac{1}{14}\begin{pmatrix} 3 \\ -2 \\ 13 \end{pmatrix}$

7.(1) $A^+=\frac{1}{42}\begin{pmatrix} 1 & 2 & 4 \\ -1 & -2 & -4 \end{pmatrix}$ (2) $A^+=\frac{1}{3}\begin{pmatrix} 2 & -3 \\ 1 & 0 \\ -1 & 3 \end{pmatrix}$

(3) $A^+=\frac{1}{8}\begin{pmatrix} 2 & -2 & 2 & 2 \\ -1 & 3 & -1 & 1 \\ 1 & -3 & 1 & -1 \end{pmatrix}$

(4) $A^{+}=\frac{1}{22}\begin{pmatrix} 2 & 8 & -2 & 4 \\ 5 & -2 & 6 & -1 \\ 5 & -2 & 6 & -1 \end{pmatrix}$

8. $X=\frac{1}{5}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

9. $X=\frac{1}{5}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$

参考文献

〔1〕 Richard Bellman, Introduction to Matrix Analysis, McGraw-Hill Book Company New York, 1970。

〔2〕 N.J.Pullman, Matrix Theory and Its Applications, Marcel Dekker, Inc. New York & Basel, 1976。

〔3〕 须田信英等，自动控制中的矩阵理论，科学出版社，1979。

〔4〕 李国平等，数学模型与工业自动控制第一卷，湖北人民出版社，1978。

〔5〕 北京大学，高等代数，人民教育出版社，1978。

〔6〕 张远达，性线代数原理，上海教育出版社，1979。

〔7〕 S.Fenyo, Modern Mathematical Methods in Technology, Vol. 2 North-Hall & Publ. 1975。

〔8〕 绪方胜彦，现代控制工程，科学出版社，1978。

〔9〕 蒋尔雄等，线性代数，人民教育出版社，1979。

〔10〕 Joel.N.Franklin, Matrix Theory, Prentice-Hall. Inc, Englewood Cliffs, New Jersey, 1963。

〔11〕 广义逆矩阵，"计算机应用与应用数学"，1974，第1期。

〔12〕 A.S.豪斯霍尔德，数值分析中的矩阵论，科学出版社，1986。

〔13〕 Alan Jennings, Matrix Computation for Engineers and Scientists, John Wiley and Sons,

London New York. Sydney. Toronto, 1977.
〔14〕N.N.汉可克，电机的矩阵分析，科学出版社，1980。
〔15〕高景德等，电机过渡过程的基本理论及分析方法上册，1982。
〔16〕何旭初等，计算数学简明教程，人民教育出版社，1980。
〔17〕蔡宣三，最优化与最优控制，清华大学出版社，1983。

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝矩阵理论及其应用
作者＝李代高编
页数＝２２８
ＳＳ号＝１０１０１７５９
出版日期＝